Missionaries, Mechanisms, and Democracy

Kevin Angell
Jeffrey J. Harden

August 2018

Users Working Paper
SERIES 2018:16
THE VARIETIES OF DEMOCRACY INSTITUTE

UNIVERSITY OF GOTHENBURG
DEPT OF POLITICAL SCIENCE

Varieties of Democracy (V-Dem) is a new approach to conceptualization and measurement of democracy. The headquarters – the V-Dem Institute – is based at the University of Gothenburg with 17 staff. The project includes a worldwide team with six Principal Investigators, 14 Project Managers, 30 Regional Managers, 170 Country Coordinators, Research Assistants, and 3,000 Country Experts. The V-Dem project is one of the largest ever social science research-oriented data collection programs.

Please address comments and/or queries for information to:

V-Dem Institute

Department of Political Science

University of Gothenburg

Sprängkullsgatan 19, PO Box 711

SE 40530 Gothenburg

Sweden

E-mail: contact@v-dem.net

V-Dem Users Working Papers are available in electronic format at www.v-dem.net.

Missionaries, Mechanisms, and Democracy*

Kevin Angell† Jeffrey J. Harden‡

July 23, 2018

Abstract

What causal pathways link Protestant missionaries to the spread of liberal democracy? Woodberry's (2012) theoretical explanation includes three central mechanisms: the development of mass printing, the expansion of education and literacy, and the growth of civil society. However, his quantitative analyses of non-Western countries focus exclusively on the positive *total effect* of conversionary Protestants (CPs) on democracy. Here, we conduct a direct empirical evaluation of the mechanisms he proposes using causal mediation methods. Our results corroborate the positive total effect, but show limited support for its causal pathways. We find minimal evidence to suggest that CPs' impact operated via mass printing or civil society. There is more support for education as a mediator, although at best it only accounted for a minority share of the total effect. We conclude that further theorizing about causal processes is necessary to strengthen the claim that Protestant missionaries contributed to democracy's rise.

Keywords: Conversionary Protestants; Liberal Democracy; Causal Mechanisms; Causal Mediation Analysis

---

*All analyses presented here were publicly documented in a preanalysis plan deposited at the Political Science Registered Studies Dataverse (https://doi.org/10.7910/DVN/SF15OP) prior to execution.

†Undergraduate Student, Department of Political Science, University of Notre Dame, 2060 Jenkins Nanovic Halls, Notre Dame, IN 46556, kangell@nd.edu.

‡Associate Professor, Department of Political Science, University of Notre Dame, 2055 Jenkins Nanovic Halls, Notre Dame, IN 46556, jeff.harden@nd.edu.

1 Introduction

In 1799, missionaries from the London Missionary Society (LMS) arrived in what is now Botswana to convert the native population to Christianity. They preached the gospel and encouraged the abolition of traditional practices such as polygamy, witchcraft, and rainmaking. However, these missionaries did more than simply proselytize their faith. They built and supported schools while greatly increasing access to books and other printed material. They also influenced politics directly, facilitating the establishment of the Bechuanaland Protectorate in an attempt to keep the country from the heavy-handed rule of the Boers. Eventually, democratic institutions strengthened in Botswana. While standard theoretical accounts might identify secular forces of modernization to explain the country's path to democracy, an alternative narrative that incorporates the legacy of conversionary Protestants (CPs)—such as those from the LMS—has recently emerged in the scholarly discussion (Woodberry 2011, 2012). Specifically, this work asks: did missionaries play a role in the rise of democracy? And if so, *how* did they exert such influence?

In a highly-cited, award-winning article, Woodberry (2012) develops a compelling theory to answer these questions. Using a discussion of the historical record as well as novel data from a sample of 142 non-Western countries, he asserts that CPs contributed to the spread of democracy via several distinct causal pathways, the most prominent being the development of mass printing, the expansion of education and literacy, and the growth of civil society. However, while he presents a considerable amount of evidence supporting a positive total effect of CPs on democracy, his quantitative analyses include no tests of these pathways. In this letter, we aim to build on Woodberry's work by conducting a direct empirical assessment of his proposed mechanisms.

Using causal mediation analysis, we recover the positive effect of CPs on democracy that Woodberry reports. However, we also find limited empirical support for the theoretical framework he proposes. Much of our evidence suggests that education mediated a noteworthy portion of the total effect, but this estimate exhibits sensitivity to specification choices. Additionally, we demonstrate little to no support for the theorized roles of mass printing or civil society. These results lead us to ultimately conclude that, despite the general robustness of the total effect, support for the claim that Protestant missionaries facilitated the rise of democracy could still be further strengthened. It is unlikely that a design-based identification strategy will ever be available to solidify the case for causality in this important research agenda. Thus, we contend that investigating evidence for proposed mechanisms is just as critical for understanding CPs' political legacy as the total effect itself. Our work here indicates that, if CPs did, in fact, contribute to the development of democracy in the non-Western world, many of the pathways by which they did so must still be established empirically.

2 Background

Woodberry (2012) advances the thesis that CPs contributed to democratic development alongside the more common scholarly explanations, such as secular rationality and economic development (244). His theoretical contention is that CPs "fostered greater separation between church and state, dispersed power, and helped create conditions under which stable democratic transitions were more likely to occur" (Woodberry 2012, 249). He proposes a comprehensive set of causal pathways to animate this theory. The most prominent of these are (1) the promotion of mass access to printed materials, (2) the spread of mass education and literacy beyond the elite classes of society, and (3) the encouragement of organizational structures as vehicles of protest, which laid the groundwork for the development of civil society. We briefly review these mechanisms here; see Woodberry (2012, 249–256) for complete details.

First, Woodberry argues that CPs greatly accelerated the growth of mass printing. CPs believed that books must be accessible to everyone so that they could easily read "God's word." Additionally, they used Protestant literature as a means of conversion, which forced other religious groups to adopt similar practices. Protestant missionaries worked to provide printed materials to the masses, which contrasted with the thinking of existing elites—that the general population was not qualified to read and interpret printed word. Woodberry notes that CPs' role in this pathway was as an initial spark. Ultimately, as printing became more widespread, market forces took control and news media that was independent of the state emerged. In short, mass printing linked CPs to democracy by paving the way for the fourth estate.

To read and understand the Bible, people needed education and literacy. Woodberry (2012) also argues that CPs catalyzed the rise of mass education around the world for this purpose (252). Specifically, he contends that CPs focused on the education of non-elite segments of society—groups that previously had little to no opportunity to attend school or learn to read. Reducing inequality in access to education subsequently expanded the group of people who were able to participate in a country's democratic development. Furthermore, he argues that economic competition ensured that this process even occurred in predominantly Catholic countries. CPs' presence spurred Catholics to provide their own mass education system, continuing the extension of schooling and literacy to non-elites (Woodberry 2012, 252).

Finally, Woodberry (2012) presents the development of civil society as a third major pathway of CPs' impact on democracy. Missionaries contributed to the dispersal of political power by facilitating the organization of opposition. They provided start-up mobilization efforts such as signature gathering for petitions and generally assisted anticolonial activists with nonviolent protest. Some of CPs' most notable legacies include publicizing colonial abuses, lobbying for policy change, and advocating for social reform in the colonies and back home.[1] These efforts helped increase individual participation in public life, even leading to the creation of political parties in some cases. Woodberry (2012) contends that this increase in political participation eventually led to the formation and success of democratic government.

Despite the richness of the theoretical framework, Woodberry's (2012) empirical analyses narrow in focus to estimation of the total effect of CPs on democracy. He employs a suite of regression analyses on a sample of 142 non-Western countries to accomplish this objective. Specifically, he models the liberal democracy measure developed by Bollen (2009), averaged over the period 1950–1994, as a function of several covariates. These include controls for alternative theories of the spread of democracy, exogenous and precolonial conditions, other factors that influenced colonizers and missionaries, and endogenous or intervening variables (see Woodberry 2012, 257–258).

[1]Woodberry (2012) separates civil society and "colonial transformation" as distinct mechanisms in his theory. We focus on civil society here as a means of simplifying our empirical analyses, but our operationalization of the concept relates to both forms (see below).

Most important, however, are the "mission variables," which represent cross-sectional country-level variation in CPs: Protestant missionaries per 10,000 population in 1923, years exposure to Protestant missions until 1960, and percent evangelized by 1900. These three "treatments" repeatedly yield regression coefficients that indicate strong positive association between CPs and democracy.

Woodberry (2012) ultimately concludes that his theory holds empirical support; by way of his proposed mechanisms, Protestant missionaries produced conditions that "laid a foundation for democracy" (268). However, this claim centers on the array of positive regression coefficients generated by the mission variables. He references some past literature that connects CPs to the various mechanisms, but never demonstrates empirically that those mechanisms do, in fact, mediate the effect of CPs on democracy. Indeed, he states that his statistical models "attempt to demonstrate a causal association between Protestant missions and democracy, but do not test which mechanism is most important" (256). In what follows we pick up the analysis at this point. We conduct a more comprehensive test of Woodberry's theory—one that involves estimation of the total effect of CPs on democracy *and* the mediating effects that he claims connects those two factors.

3 Research Design

We conduct an empirical test of Woodberry's (2012) three primary mechanisms with the methodology for causal mediation analysis developed by Imai, Keele, Tingley, and Yamamoto (2011).[2] This approach, which we discuss in more detail in the appendix, quantifies a causal mechanism with a series of two regression models: one that regresses a *mediator*—a quantitative measure of the mechanism—on a treatment variable and covariates, then a second that regresses the outcome on the treatment, mediator, and covariates. The algorithm generates predicted values of the mediator by manipulating treatment status in the mediator model. Those two values are then entered into the mediator variable in the second model to produce potential outcome predictions.

The difference in the predictions from the outcome model represents a key quantity of interest:

[2] Mediation analysis has existed in many forms for decades. We employ Imai et al.'s (2011) implementation because it rigorously connects the role of mechanisms to the potential outcomes framework of causal inference.

the *average causal mediation effect* (ACME), which provides an empirical estimate of the (average) strength of the mechanism. A typical approach to quantifying the uncertainty of this estimate, which we follow here, is to bootstrap the entire process (Imai et al. 2011). Another relevant quantity is the average direct effect (ADE), or the portion of the total effect that does *not* operate through that particular mechanism. The total effect is represented by the sum of the ACME and ADE. Thus, a useful means of substantively evaluating an ACME estimate is to consider the proportion of the total effect mediated by the mechanism of interest, computed by dividing the estimated ACME by the estimated total effect.

3.1 Model Specification

To employ this methodology, we must select a model specification. Woodberry (2012) presents more than 30 distinct regressions in the main text of his article and dozens more in the supplementary materials. Thus, our own analyses carry high risk for the problems associated with "researcher degrees of freedom," such as emphasizing only those specifications that conform to a preconceived set of expectations (explicitly stated or not). To reduce this risk, we deposited a preanalysis plan for this research at the Political Science Registered Studies Dataverse in June 2018, prior to observing any mediation analysis results (see the appendix for the complete text). In that document we selected Woodberry's (2012) Model 4 from Table 3 (262) as our main model specification due to its comprehensive coverage of the theoretical framework. We also preregistered several alternative specifications that we thought merited consideration; we present some of those results below and the remainder in the appendix.

We replicated Woodberry's (2012) robust regression estimates exactly, then implemented mediation analysis using Protestant missionaries in 1923 as our treatment variable. We justified this choice over the other two mission variables in our preanalysis plan. Briefly, the years exposure variable measures the *timing* of missionaries, not the more conceptually-accurate missionary levels. Additionally, the percent evangelized measure is problematic because it includes both Protestant *and* Catholic conversions (Woodberry 2012, 257).[3]

[3] See the appendix for results with each of the other two mission variables as treatment.

3.2 Measuring Mechanisms

A key feature of our analysis is the addition of measures of the mechanisms to the regression models. We employ indicators from Woodberry's (2012) own data as well as from additional sources. Our data collection efforts yielded multiple candidate measures. We report results with all of them, but in choosing a set on which to focus we considered the following characteristics, in order of importance: (1) conceptual match with Woodberry's theoretical framework, (2) data originating in Woodberry's replication materials, and (3) amount of missing data. Regarding this third point, all of the measures have some missingness. Thus, we employed multiple imputation with the Amelia II software (Honaker, King, and Blackwell 2011) to maintain the original sample.[4]

We measure mass printing with Woodberry's (2012) data on average daily newspaper circulation per 1,000 population. This indicator aligns with his contention that CPs' use of printed material gave rise to a robust news media. For mass education, we use a Gini coefficient measure of education inequality from the Varieties of Democracy Project (V-Dem, see Coppedge et al. 2018*a*). This variable closely matches Woodberry's theoretical discussion of CPs' influence on education among non-elites (see above). Finally, we employ V-Dem's civil society participation index for the third mediator. This index measures societal involvement in civil society organizations (CSOs), whether major CSOs are consulted by policymakers and thus involved in governance, whether women participate in CSO, and whether nominations of candidates within political parties are decentralized.

We provide detailed discussions of our measurement choices in the preanalysis plan and appendix. However, one point that is important to note before proceeding is the issue of timing. Woodberry's (2012) models are cross-sectional in nature. The treatment variable represents the state of CPs in 1923 and the outcome is an average democracy score over the period 1950–1994. We also average over several years' worth of data with our mediators, which leads to the question of which timeframe to choose. In our main analyses below, the newspaper circulation variable is

[4]See the appendix for details on combining multiple imputation with mediation analysis, diagnostic reports on the quality of the imputations, and results from addressing missing data with listwise deletion.

averaged over the earliest years available in Woodberry's data (1975, 1980, 1985, and 1990) and the V-Dem measures are averaged over 1924–1994. Thus, these measures occur contemporaneously with the outcome. This approach is somewhat at odds with the concept of a mediator, which should come between treatment and outcome in the causal sequence (Imai et al. 2011).

Accordingly, we also consider versions of the V-Dem measures, as well as a new mass printing measure, averaged over 1924–1949.[5] Syncing the temporal sequence is conceptually helpful, but also places more burden on the imputation procedure because larger proportions of the mediation measures are missing for earlier years. Admittedly, we did not take a firm stand in our preanalysis plan on which strategy to present in the main analyses here. We ultimately decided (post-hoc) to choose the versions measured partially concurrent with the outcome for the sake of using less imputed data, but we report results with the other approach in the appendix.

4 Results

We begin by estimating the mediation effects of newspaper circulation, education inequality, and the civil society participation index using the main model specification discussed above. Figure 1 presents the ACMEs, ADEs, and total effects for each of these mediators. We generate these estimates for each mediator sequentially—ignoring the other two variables at each iteration—which produces slight variation in the reported total effects due to bootstrapping error. The more important point to note is that our analysis recovers (within bootstrapping error) the same total effect that Woodberry (2012) reports: an estimate of 4.43 with a 95% confidence interval of (1.28, 7.59). Substantively, a standard deviation increase in the treatment variable corresponds with a 0.26 standard deviation increase in the outcome.

[Insert Figure 1 here]

Moving to the mediation results, the graph shows that none of the ACMEs or ADEs are statistically significantly different from one another. This lack of certainty is due, in part, to the relatively small sample of data as well as our use of multiple imputation. There are limits to what we can

[5] Specifically, we switch the newspaper circulation measure with Fink-Jensen's (2015) measure of book titles published per capita.

learn from 142 countries with some missing data. Nonetheless, with appropriate caution we contend that assessing point estimate magnitude as well as testing for individual null effects can still be useful and informative.

The ACME of 0.57 for our mass printing mediator is fairly small, although not completely negligible. It comprises about 13% of the total effect. The civil society ACME is essentially zero (0.06), mediating just 1% of the total effect. Moreover, neither of these estimates are statistically significant at the 0.05 level. In contrast, the education estimates in Figure 1 clearly demonstrate a mediation effect. The ACME is 1.53, which represents 35% of the total effect. Although this estimate is not statistically distinguishable from the other ACMEs, its 95% confidence interval is bounded away from zero. In short, a model specification that we chose *a priori* for its prominence in Woodberry's (2012) analysis demonstrates favorable evidence for just one of three causal mechanisms.

4.1 Alternative Specifications

While the main model is comprehensive and theoretically-informed, it is only one of many models that could be used to test the theory. Thus, we also must consider the robustness of our results in Figure 1 to other reasonable specifications.[6] As with our choice of the main model, we discussed several plausible alternatives in our preanalysis plan, then executed them. We briefly outline these alternatives here; see the preanalysis plan and appendix for details.

First, we add controls for settler mortality and gross domestic product (GDP) per capita to the main specification. Past work points to these variables as important alternate explanations of the rise and spread of democracy (e.g., Acemoglu, Johnson, and Robinson 2001; Woodberry 2012, 263). Our second alternative specification uses the main model with different measurement strategies for the mass printing and mass education mediators: book titles per capita (Fink-Jensen

[6] We also estimated the main model specification with *multiple mediation analysis*. This approach relaxes a key assumption of standard mediation analysis: no association between mediators (see the appendix). However, it requires that we use a binary version of the treatment variable (Imai and Yamamoto 2013). Thus, this approach cannot provide much insight into the question of the mediators' absolute roles as causal mechanisms because the treatment variable is considerably different. It is helpful in checking the robustness of their relative magnitudes, and we find that our conclusions are unaffected in this regard.

2015) for the former and the literate proportion of the population for the latter (Coppedge et al. 2018*a*). Third, we estimate the effects using another specification that Woodberry (2012) features prominently: Table 2, Model 3 (260). This model omits covariates related to the "process of colonization." Finally, we consider results with our own "preferred" model specification, which addresses several issues that we noted as we replicated Woodberry's original results. Specifically, our preferred model avoids posttreatment bias, streamlines the definition of treatment, and conserves degrees of freedom.

We summarize the essential results of these alternative specifications here (see the appendix for further details). Specifically, Table 1 reports, for each mediator and each alternative specification, (1) the proportion of the total effect mediated and (2) whether the corresponding ACME is statistically significantly different from zero.

[Insert Table 1 here]

Several patterns stand out in Table 1. First, all of the specifications yield the same relative ordering with respect to proportion mediated. Mass education is the strongest mediator, followed by mass printing, then civil society. The magnitudes of these values shows some heterogeneity; in some cases education is clearly the strongest while in others its effect is essentially equivalent to that of mass printing. But the weight of the evidence indicates that education is the most important of the three mechanisms. The fact that mass education is the only mediator that reaches statistical significance in *any* specification underscores this point. Indeed, even in cases in which the substantive magnitude of the mass printing mediator increases to noteworthy levels (e.g., specifications #2 or #4), the uncertainty surrounding those estimates tempers the inferences we could draw. Finally, Table 1 clearly indicates that there are other pathways that transmit the effect of CPs on democracy. Summing the proportions mediated within each specification still leaves an average of 60% of the total effect in the ADEs, which strongly suggests that additional mechanisms and/or a direct effect characterize the total causal process.

5 Conclusions

Understanding causal mechanisms is often considered a fundamental component of social science (e.g., Deaton 2010; Imai et al. 2011, but see Holland 1988). Studying pathways and processes draws the focus away from causality as a "black box" and emphasizes the theoretical heart of a substantive research question (Hedström 2008; Imai et al. 2011). Woodberry's (2012) study of the impact of Protestant missionaries on democracy is a compelling example of this perspective. The causal relationship he proposes may be unintuitive to many given that the scholarship on which he builds largely ignores the role of activist religion in explaining democratic development (Woodberry 2012, 244). We contend that a crucial aspect of confronting such skepticism involves developing mechanisms in a theoretical framework, as Woodberry does, then *testing* those mechanisms empirically.

We build on Woodberry's (2012) important work by completing the second of those two tasks. Using causal mediation analysis, we empirically assess the role of the three central causal pathways that he proposes to explain the positive influence of CPs on the spread of democracy: the development of mass printing, the expansion of education and literacy, and the growth of civil society. We confirm the overall positive effect of CPs, but find limited support for the theory itself. The results weakly suggest that mass printing may have mediated a small portion of the effect, demonstrate virtually no support for civil society's effect, and yield stronger, but not entirely robust, evidence in favor of education as a mediator. Ultimately, we conclude that the nature of Protestant missionaries' political legacy in the non-Western world remains open to some degree. Woodberry's (2012) theory is well developed, but empirically it explains only a minor share of the total effect. In sum, more theoretical appraisal of the causal processes involved is necessary to determine how spreading the Gospel helped spread democracy.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development." *American Economic Review* 91(5): 1369–1401.

Bollen, Kenneth A. 2009. "Liberal Democracy Series I, 1972–1988." *Electoral Studies* 28(3): 368–374.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jerey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig, and Daniel Ziblatt. 2018*a*. "V-Dem Codebook v8." Varieties of Democracy (V-Dem) Project. https://www.v-dem.net/.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jerey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig, and Daniel Ziblatt. 2018*b*. "V-Dem [Country-Year/Country-Date] Dataset v8." Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemcy18.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48(2): 424–455.

Fink-Jensen, Jonathan. 2015. "Book Titles per Capita." IISH Dataverse. http://hdl.handle.net/10622/AOQMAZ.

Hedström, Peter. 2008. "Studying Mechanisms to Strengthen Causal Inferences in Quantitative Research." In *The Oxford Handbook of Political Methodology*, ed. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier. New York: Oxford University Press.

Holland, Paul W. 1988. "Causal Mechanism or Causal Effect: Which is Best for Statistical Science?" *Statistical Science* 3(2): 186–188.

Honaker, James, Gary King, and Matthew Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45(7): 1–47.

Imai, Kosuke, and Teppei Yamamoto. 2013. "Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments." *Political Analysis* 21(2): 141–171.

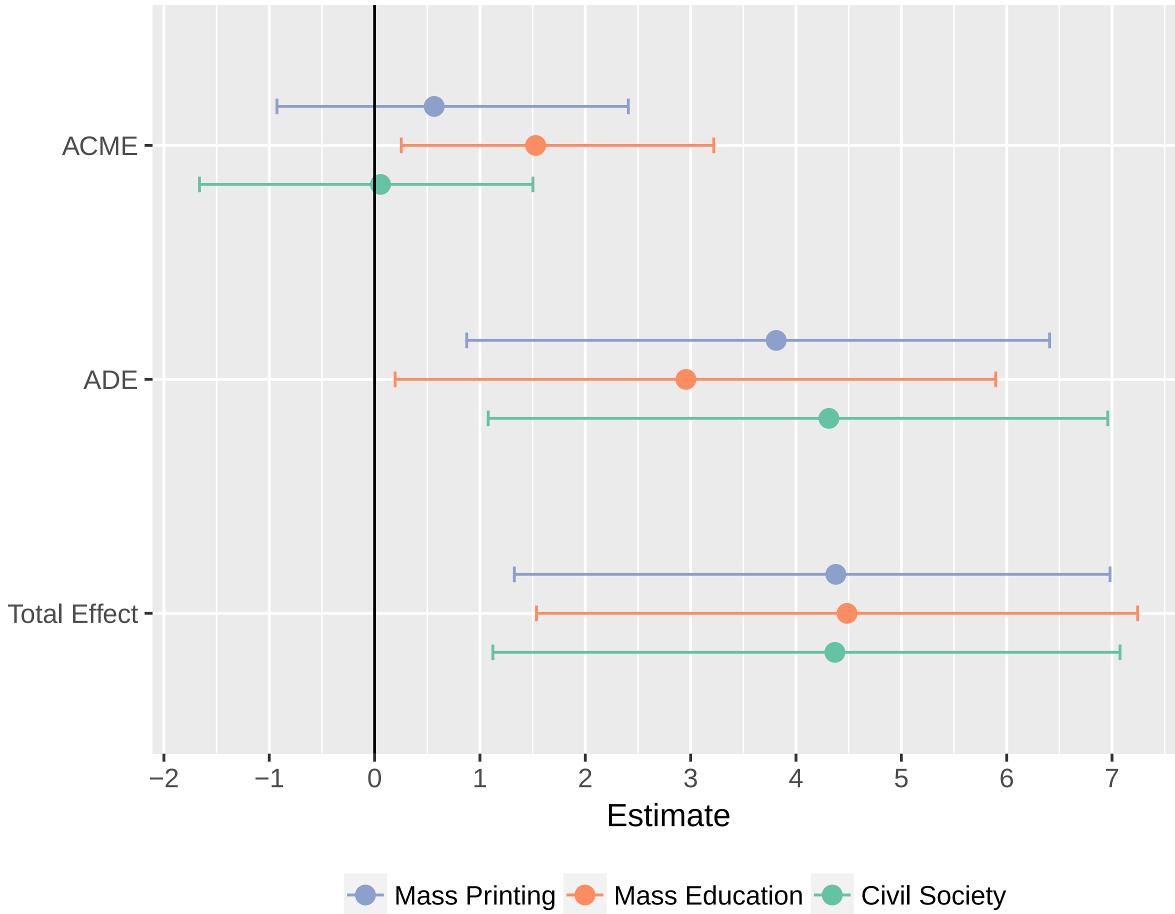
Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4): 765–789.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi ting Wang, Joshua Krusell, and Farhad Miri. 2018. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21, 3d edition.

Woodberry, Robert D. 2011. "Ignoring the Obvious: What Explains Botswana's Exceptional Democratic and Economic Performance in Sub-Saharan Africa?" Project on Religion and Economic Change Working Paper #05. https://doi.org/10.13140/RG.2.1.3027.4641.

Woodberry, Robert D. 2012. "The Missionary Roots of Liberal Democracy." *American Political Science Review* 106(2): 244–274.

Figure 1: The Mediation Effects of Mass Printing, Mass Education, and Civil Society



Note: The graph presents the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals.

Table 1: Summary of Mediation Results with Alternative Specifications

Specification		Mediator	% Mediated	Significance
1	Original Results (Figure 1)	Mass Printing	13	–
		Mass Education	35	✓
		Civil Society	1	–
2	Control for Settler Mortality and ln(GDP per Capita)	Mass Printing	18	–
		Mass Education	30	✓
		Civil Society	4	–
3	Alternative Printing and Education Measures	Mass Printing	13	–
		Mass Education	14	–
		Civil Society[a]	–	–
4	Table 2, Model 3 (Woodberry 2012, 260)	Mass Printing	22	–
		Mass Education	23	✓
		Civil Society	6	–
5	Authors' Preferred Specification	Mass Printing	4	–
		Mass Education	13	–
		Civil Society	2	–

Note: Cell entries report the proportion of the total effect mediated and whether the corresponding ACME is statistically significantly different from zero (denoted by ✓) or not (–). [a] Civil society results are not reported in specification #3 because there is no alternative measure for that mediator.

Missionaries, Mechanisms, and Democracy

Supplemental Appendix

Kevin Angell* Jeffrey J. Harden†

July 23, 2018

Abstract

This appendix contains supplementary information, robustness checks, and diagnostics for the analyses described in the main text.

Contents

*Undergraduate Student, Department of Political Science, University of Notre Dame, 2060 Jenkins Nanovic Halls, Notre Dame, IN 46556, kangell@nd.edu.

†Associate Professor, Department of Political Science, University of Notre Dame, 2055 Jenkins Nanovic Halls, Notre Dame, IN 46556, jeff.harden@nd.edu.

1 Summary

This appendix presents additional information intended to provide a complete picture of the analyses we conducted to answer our research question: what causal pathways link Protestant missionaries to the spread of liberal democracy? In undertaking this project we were keenly aware that a key challenge would be the availability of many potential researcher degrees of freedom. Measurement, modeling strategy, subsetting, interpretation, and other aspects of the research process provide many reasonable choices in most quantitative analyses, and this one is no exception. Indeed, the supplementary information for Woodberry's (2012) original article is 192 pages. As we note in the main text, our primary strategy for addressing this issue was to document our choices in a preanalysis plan. All of the analyses presented in the main text and this document were preregistered and justified prior to observing results. This approach does not eliminate the possibility that our own biases influenced our analyses in some way, but it does reduce the threat.

In the main text, we conclude that there is "limited" empirical support for Woodberry's (2012) three key causal mechanisms. We base this claim on the combined assessment of the substantive magnitude and uncertainty of the estimated ACMEs for mass printing, mass education, and civil society. The mass printing ACME shows reasonable substantive magnitude in some specifications, but its confidence interval is never bounded away from zero. The mass education ACME is more promising; in several specifications it is substantively larger than the others and statistically significant. But in others—including our preferred specification—it yields a substantively smaller and statistically nonsignificant mediation effect. Finally, civil society consistently produces ACMEs that are substantively small (near zero) and statistically nonsignificant.

Broadly speaking, the analyses presented in this appendix reflect the results from the main text: education almost always produces the strongest mediation effect, but also displays sensitivity to several analytic choices. Additionally, while the other two effects sometimes increase in magnitude compared to what we report in the main text, they are *never* statistically significant across all of the analytic strategies we employed. Thus, we place the most confidence in the conclusions drawn from the results in the main text. While the analyses we present in this appendix demonstrate some heterogeneity, we contend that the main text results represent the best of, admittedly, many possible permutations of these analyses.

2 Causal Mediation Analysis

In this research, we employ causal mediation analysis to empirically assess the causal mechanisms proposed in Woodberry (2012). Here, we provide some details of the methodology, its assumptions, and our implementation of it. This summary is intended to be a starting point for readers unfamiliar with these methods. It is not a comprehensive discussion.

The core motivation for causal mediation analysis is that social scientists are not only interested in whether one factor causes another or even by how much, but also *why* that relationship appears. Methods for assessing mechanisms empirically have been used for several decades (e.g., Haavelmo 1943; Baron and Kenny 1986). More recently, Imai, Keele, Tingley, and Yamamoto (2011) have updated and expanded mediation analysis to align with the potential outcomes framework of causal inference (see also Imai, Keele, and Yamamoto 2010; Imai and Yamamoto 2013; Keele, Tingley, and Yamamoto 2015). We use the R package `mediation` to implement this approach in our analyses here (Tingley, Yamamoto, Hirose, Keele, and Imai 2014).

2.1 Assumptions

Under a set of assumptions, the Imai et al. (2011) methodology yields nonparametric identification of the ACME for a proposed mediator. First, we assume sequential ignorability, which is essentially two assumptions made in order (Imai et al. 2011, 770). We assume that treatment assignment is independent of potential outcomes and potential mediators, conditional on the pretreatment covariates. Additionally, we assume that the mediator is independent of potential outcomes, given treatment status and pretreatment covariates. These are potentially strong assumptions that are not easily testable. However, Imai et al. (2011) provide an approach to sensitivity analysis that allows the analyst to assess robustness of the estimated effects to possible violations of sequential ignorability (see section 6 below).

Second, "standard" mediation analysis assumes that the mediator is independent of alternative mediators. Imai and Yamamoto (2013) propose a methodology for relaxing this assumption, which we also employ here. In mediation analysis with multiple mediators, we need only assume that the main mediator of interest is independent of potential outcomes after conditioning on an alternative mediator, the treatment, and pretreatment covariates. This approach is useful here because we have three mediators of substantive interest. However, it is limiting in that we can only control for one alternative at a time. Additionally, the current implementation is only derived for binary treatment variables (Imai and Yamamoto 2013).

Finally, when conducting mediation analysis with multiple mediators we assume no interaction between the treatment and mediator (Imai and Yamamoto 2013, 157). This assumption is also quite strong, and so we use Imai and Yamamoto's (2013) methodology for testing for interactions as well as for conducting sensitivity analyses to potential violations (see section 6 below).

2.2 Estimation

Estimation is done in several steps (see Imai et al. 2011, 773–774). First, the analyst models the mediator as a function of the treatment variable and pretreatment covariates. Second, he or she models the outcome variable as a function of the mediator, treatment, and the same covariates. Next, the estimated mediator model is used to predict two values of the mediator: one under control and one under treatment. These values are then entered into the outcome model to generate outcome predictions, and the difference in those predictions is the ACME estimate. To obtain measures of uncertainty, this entire process can be repeated in a bootstrapping procedure.

In our implementation of the method we estimate the ACMEs with the process described above and compute 95% confidence intervals for our estimates using the 2.5 and 97.5 percentiles of the ACME distributions. Additionally, because we must impute some missing data, we repeat the entire process for each estimate 10 times using 10 imputed datasets. We bootstrap 100 replicates with each dataset, then combine them to form 1,000 replicates in a single bootstrap distribution. This approach accounts for the added uncertainty that comes from imputing data (Blackwell, Honaker, and King 2017, 309).

3 Mediator Data

Here, we fully describe our choice of data to measure the mechanisms of interest. Recall that our priorities in selecting data were (1) conceptual match with Woodberry's theoretical framework, (2) data originating in Woodberry's replication materials, and (3) amount of missing data.

Beginning with mass printing, we considered two sources: data from Woodberry (2012) on newspaper circulation and data from Fink-Jensen (2015) on the number of books titles published per capita. Woodberry's (2012) data include the average daily newspaper circulation per 1,000 population in 1975, 1980, 1985, and 1990.[1] We compute the means across these four years to construct a mass printing mediator variable. The measure itself is internally valid—it aligns with the contention that CPs' use of printed material gave rise to a robust news media. However, the timing is not ideal. Measurement does occur posttreatment (i.e., after 1923), but it also falls after the measurement of the outcome begins (1950). Fink-Jensen's (2015) number of book titles per capita measure allows us to obtain data that is both posttreatment and premeasurement of our outcome (1924-1949).[2] However, this source have a much larger amount of missingness for Woodberry's sample of countries than the newspaper data. Thus, to avoid the extremely large volume of imputation needed to use book titles per capita, we rely on Woodberry's (2012) newspaper data.

The next mediator is education. Woodberry's data also contain a candidate measure, which he uses in some robustness checks (Woodberry 2012, Table 5, 265). Specifically, his data include the mean enrollment in secondary education from 1960–1985 (Barro and Lee 1994). However, this variable does not cover all of the countries in his full sample and is also partially concurrent with the democracy outcome. Additionally, mean enrollment in secondary education does not completely capture the impact of CPs via education as proposed in Woodberry (2012). He argues that CPs spread literacy and education to non-elites, neither of which would necessarily appear as an increase in secondary school enrollment because the development of literacy occurs in primary school and the expansion of education could reflect the spread of primary schooling. As an alterna-

[1] These variables come from the United Nations (UN) Data (see http://data.un.org/).

[2] This measure is available at https://www.clio-infra.eu/Indicators/BookTitlesperCapita.html.

tive, we gathered education data from the Varieties of Democracy Project (V-Dem, see Coppedge et al. 2018*a*). V-Dem provides data on the percentage of a country that is literate (from Vanhanen 2003), the average years of education for citizens older than 15, and the Gini measure of education inequality that we argue best captures Woodberry's theoretical claim. Each of these variables are available at least as early as 1924.

Finally, civil society is the third causal mechanism we consider. Although the development and spread of voluntary organizations, nonviolent protest, and other political movements is a key element of Woodberry's (2012) theoretical framework (e.g., 252–253), he does not include any such measure in his analyses or replication data. To obtain one, we again turn to V-Dem (Coppedge et al. 2018*a*). V-Dem provides a civil society participation index that captures the extent to which citizens are involved in CSOs, how much CSOs are consulted by policy makers, whether women participate in CSOs, and whether political candidate nomination is decentralized (Coppedge et al. 2018*a*). This variable is highly suitable to serve as a mediator because it provides a robust measurement of the civil society concept that Woodberry claims connects CPs with the development of democracy. It also covers the vast majority of the countries in his estimation sample (Woodberry 2012, 252–253).

For each of these sources of data, a key issue that we must address is missing data. Our objective is to use Woodberry's full sample of data in our mediation analyses, including some countries that are not covered in our data on mediators. In his analysis, Woodberry's (2012) measurement strategy involves computing several variables by averaging over a span of years; for instance, his outcome variable is a mean level of democracy over the period 1950–1994. We adopt this approach in constructing our mediators, which helps reduce the impact of missing data. However, the historical nature of the data means that some countries did not exist or were known by different names when our various mediators were measured.[3] Consequently, we consider two options when constructing mediators: compute the averages from (1) 1924–1949 or (2) 1924–1994. The former approach best matches the temporal nature of a mediator because it occurs between treatment

[3]See Woodberry (2012, 257) for a discussion of how he addresses this issue in measuring the outcome variable.

(1923) and outcome (1950). However, it yields more missing data. In contrast, the second strategy overlaps measurement of the outcome, but allows for more data to be collected. We conduct our analyses using both measurement strategies to assess the robustness of our results. Table A1 summarizes our mediator data collection efforts.

For those data values we cannot fill in, we use multiple imputation with the Amelia II software available in R (Honaker, King, and Blackwell 2011). Amelia begins with the assumption that the complete data are distributed multivariate normal, then employs the expectation-maximization (EM) algorithm to impute the missing values. It repeats this process m times (we set m to 10), then analysis continues as usual with those m complete datasets. An adjustment to measures of uncertainty is also necessary (see section 2.2). We use the complete data from Woodberry's sample to impute the missing values in our mediators.[4]

[Insert Table A1 here]

4 Alternative Model Specification Results

Here we fully describe the results from the alternative specifications of standard mediation analysis that we report in the main text (Table 1). Recall that we first outlined these alternatives in our preanalysis plan to guard against the problems associated with researcher degrees of freedom. The graphs in Figure A1 report the same information as Figure 1 of the main text using four specifications: (1) the main model specification with controls for settler mortality and gross domestic product (GDP) per capita, (2) the main model with alternative measurement strategies for the mass printing and mass education mediators, (3) a simpler specification from Woodberry's results, and (4) our own "preferred" model specification, which addresses several issues that we noted while replicating the original results. See Table A4 in section 7.1 for complete lists of the variables appearing in each specification.

[Insert Figure A1 here]

[4] We also impute missing values in two additional covariates: settler mortality and GDP (see section 9).

We begin by adding two important covariates from Woodberry's data to the main model. The results in the top panel of Figure A1 come from a specification that includes controls for European settler mortality (Acemoglu, Johnson, and Robinson 2001) and economic development, measured as the natural log of GDP per capita (averaged over 1960–1994). Woodberry (2012) controls for both of these variables in some specifications, but not the main model. We include them here because they represent theoretically-informed alternate explanations for the rise and spread of democracy (for summaries, see Woodberry 2012, 263).[5]

Empirically, including settler mortality and GDP yields fairly minor changes to our results. The total effect weakens slightly, as does the ACME for education, but both of those estimates remain statistically significant. Substantively, this specification indicates that education inequality mediates about 30% of CPs' effect on democracy. In contrast, the ACMEs for the other mediators *increase* in magnitude. Newspaper circulation mediates 18% of the effect and the civil society participation index mediates 4%. However, the 95% confidence intervals for these latter two effects both include zero.

Our next analysis returns to the main model specification, but alters the measurement strategy for mass printing and mass education. We measure the former using book titles per capita (Fink-Jensen 2015) and the latter as V-Dem's indicator for the literate percentage of the population. We average both variables over the period 1924–1994. The book titles measure is conceptually similar to Woodberry's newspaper circulation variable, but much of the data are missing (79%). Literacy rate is somewhat different from the education inequality measure, although it captures a different aspect of the original theory. Instead of focusing on the proposed pathway in which CPs facilitated education for the non-elite masses, literacy taps into CPs' objective of giving everyone the ability to read the Bible. Note that the civil society measure is unchanged here, so its specification is identical to our main model (although the estimates are slightly different due to bootstrapping error).

The top right panel of Figure A1 indicates that these changes are consequential for results. The

[5]Both settler mortality and GDP per capita require multiple imputation of missing data. See section 9 for details.

book titles measure produces essentially the same ACME point estimate as in Figure 1 of the main text (13% mediated). However, its confidence interval is quite large due to the fact that so much of the variable must be imputed. Measuring mass education as percent literate produces a notable drop in the magnitude of that mediator, down to 14% of the total effect. Furthermore, the education ACME is not statistically significant in this specification.

In the bottom left panel of Figure A1 we return to our original mediators and estimate the effects with a simpler model specification that Woodberry (2012) reports: Table 2, Model 3 (260). The critical difference between this specification and our main model is that the former does not include covariates representing the "process of colonization," including ease of access and perceived value of each country as a colony (Woodberry 2012, 262–263). The graph indicates that this change yields somewhat different results from Figure 1 of the main text. The total effect weakens to about 3.75, though it remains statistically significant. Additionally, the ACMEs for mass printing (22% mediated) and mass education (23% mediated) are essentially the same, although only the confidence intervals of the latter are bounded away from zero. Civil society again produces a small, nonsignificant mediation effect (6%).

Finally, we consider our own preferred model. As our preanalysis plan describes, this specification is intended to address several issues that arose as we replicated the original results. First, we removed all of the covariates that are measured after 1923 to avoid posttreatment bias (Montgomery, Nyhan, and Torres 2018).[6] We include settler mortality—which is measured pretreatment—due to its theoretical relevance. We also include indicator variables for five regions of the world, as defined by Woodberry.[7] We also reduce down to a single treatment variable for the sake of definition clarity: Protestant missionaries per 10,000 population in 1923. We omit percent evangelized by 1900 because it includes Catholic and Protestant evangelization and years exposure to Protestant missions because it is partially posttreatment. Finally, we simplify the model by removing several

[6]These include percent European and percent Muslim, as well as the additional covariate of GDP per capita.

[7]These regions are Sub-Saharan Africa, Asia, Latin America and the Caribbean, the Middle East/North Africa, and Oceania (the reference category). Our original plan was to interact these region indicators with the treatment variable to assess variation in the treatment effect by region. However, upon executing that plan we discovered that the data were not up to the task (see section 8.4).

variables associated with the perceived value of a country. While Woodberry (2012) provides theory for these variables, only one—an indicator for whether a Protestant colonizer took a colony from a Catholic colonizer—yields a substantively and statistically significant coefficient estimate. Thus, to save degrees of freedom, we only retain that variable in our preferred specification.

Results from our preferred specification appear in the bottom right panel of Figure A1. We note that the confidence intervals are somewhat smaller in that graph, suggesting that our specification choices may have improved efficiency. The total effect is slightly weaker than what we show in Figure 1 of the main text, but remains statistically significant. The mediation effects, however, all move toward zero compared to the previous results. The proportions of the total effect mediated are 4% (mass printing), 13% (mass education), and 2% (civil society). All of these estimates are statistically nonsignificant. In other words, the relative ordering of the mediation effects remains the same, but in absolute terms this specification suggests that the three mechanisms explain very little of the CPs' impact on democracy.

5 Controlling for Association Between Mediators

The analyses we present in the main text assume that there is no association between mediators. This assumption could be problematic, and thus a potentially better approach would be to relax that assumption and control for the alternative mediators. We do so here using the methodology proposed by Imai and Yamamoto (2013).

5.1 Empirical Assessment

First, we empirically assess the association between mediators. Imai and Yamamoto (2013) recommend conducting this assessment by regressing one mediator on alternative mediators, treatment, and the covariates (167). We report these as least squares regressions for each mediator/alternative combination in Table A2. We use the main model specification, although we only present the alternative mediator coefficients and standard errors to conserve space (see the replication materials for full results).

[Insert Table A2 here]

The results in Table A2 indicate that there is statistically significant association between mass printing (newspaper circulation) and mass education (education inequality) as well as mass printing and civil society (participation index). The association between education and civil society is not statistically significant. In short, relaxing the assumption of independent mediators is worthwhile for most of our mediation estimates.

5.2 Estimation and Results

To estimate the mediation effects controlling for alternative mediators we must make an adjustment to our treatment variable. The methodology proposed by Imai and Yamamoto (2013) is only derived for binary treatment variables, and thus we must recode our treatment accordingly. Our objective in recoding was to generate a binary treatment that produced a total effect similar to the main model's original estimate with respect to magnitude and statistical significance. We made this decision in an attempt to maintain as much comparability with the results presented in the main text as possible.

We ultimately chose the 75^{th} percentile of the original treatment variable as the cutoff: cases above this threshold were coded as treated and all others were coded as untreated. This decision yields a total effect estimate of 14.37 with a 95% confidence interval of (4.63, 24.11). Thus, the total effect retains statistical significance in this analysis, but the magnitude is larger. Other logical thresholds, such as the median, produced nonsignificant total effect estimates near zero. This result suggests that much of the total effect reported in Woodberry (2012) is driven by cases with very large values of the treatment variable (i.e., countries with large numbers of CPs relative to the population size).

Additionally, the multiple mediator estimation routine allows for one main mediator and one alternative mediator. Because we have three mediators in total, we conduct the analysis twice for each mediator to estimate the effects after controlling for each of the other two. We present the results in Figure A2.

[Insert Figure A2 here]

The top left panel of Figure A2 shows the mediation effects using our binary treatment measure with *standard* mediation analysis (i.e., assuming independence between mediators). We show these results to confirm that our recoding to a binary treatment recovers the same relative ordering of mediation effects as in the analyses with the original treatment variable. Of course, the scale has also changed because the total effect is larger. However, the ACMEs are even somewhat larger in proportion to the total effect compared to what we report in the main text: 2.87 (21% mediated) for newspaper circulation, 5.72 (41%) for education inequality, and 2.07 (15%) for civil society. However, as in our main results the three ACMEs are not statistically significantly different from one another and the education ACME is the only one that is significantly different from zero.

We present the mediation effects of education inequality and civil society controlling for newspaper circulation in the top right panel. The graph shows very little change to the estimates. The lower confidence bound on the ACME for education inequality just crosses zero after controlling for newspaper circulation. But the magnitude of the estimate remains similar to what we report here and in the main text (36% mediated). The ACME for the civil society participation index is essentially unchanged compared to its estimate in the top left graph.

Next, the bottom left panel of Figure A2 gives the results for newspaper circulation and civil society participation index controlling for education inequality. The ACME estimate for the former shows no change from the analogous estimate in the top left panel. The civil society estimate increases in magnitude to 25% mediated, although it remains statistically nonsignificant. Finally, the bottom right panel presents the estimates for newspaper circulation and education inequality after controlling for civil society. Both ACMEs remain similar to their previous values from the top left panel. The education inequality estimate changes the most, increasing in magnitude to 45% of the total effect.

Overall, the broad trend of this analysis conforms to what we report in the main text. Relaxing the assumption of independence between mediators does not produce dramatic shifts in our substantive conclusions regarding the relative importance of the mediators. The scale of the effects increases in magnitude in these results. However, the top left panel of Figure A2 suggests that those increases are mostly a product of incorporating a binary version of the treatment variable, not controlling for alternative mediators. This pattern is important to note because the ACMEs for mass printing and civil society move away from zero in these analyses, which contrasts with the results presented in the main text. However, we regard these changes largely as statistical artifacts of a suboptimal, but necessary, choice. Moreover, consistent with our main analyses neither of those estimates ever reaches statistical significance and the relative ordering of the mediators' effects remain consistent with what we report elsewhere in this research.

6 Sensitivity Analyses

A key advantage of Imai et al.'s (2011) framework for causal mediation is the availability of sensitivity analyses to various assumptions. Here we assess the sensitivity of our results to confounding from an omitted pretreatment covariate as well as the possibility of treatment-mediator interactions.

6.1 Sensitivity to Pretreatment Covariates

We begin by assessing the sensitivity of the standard mediation estimates reported in Figures 1 of the main text and Figure A1 to hidden confounding from pretreatment covariates. Imai et al. (2011) note that the hypothetical confounding influence of an omitted variable can be parameterized as correlation between the error terms of the mediator and outcome models, which they denote ρ (see also Tingley et al. 2014, 13–16). The graphs in Figures A3–A7 plot ρ on the x-axes against the ACME and ADE estimates (and shaded 95% confidence intervals) on the y-axes for each mediator. The dashed lines mark the original estimates, in which ρ is assumed to be zero (i.e., no omitted variable).[8] The ideal results would be a nearly flat line, indicating little movement in the estimated effect, even in the presence of confounding. In practice, such a finding does not always appear, in which case it is important to consider at what values of ρ the estimate changes sign or becomes statistically nonsignificant.

We begin with Figure A3, which shows sensitivity results for the main analyses presented in

[8] The dashed lines do not fall exactly on the estimated effects reported in the main text due to random bootstrapping error.

Figure 1 of the main text.[9] Those graphs suggest that there is little sensitivity in the effects for mass printing and civil society. However, the ACME and ADE estimates for mass education display more sensitivity. If an omitted variable produces a negative correlation of just -0.15 between the two error terms exists, the ACME becomes negative.

[Insert Figure A3 here]

Next, Figure A4 shows sensitivity results for the analyses that include controls for settler mortality and GDP. The graphs again suggest that there is little sensitivity in the effects for mass printing and civil society. The ACME and ADE estimates display less sensitivity compared to the main model results. In this case, the ACME does not become negative until ρ reaches -0.35.

[Insert Figure A4 here]

Figure A5 shows sensitivity results for the analyses with the alternative printing and education measures. The estimates for mass printing show more sensitivity with the book titles measure; at ρ = 0.30 the ACME becomes negative. The mass education ACME becomes negative at $\rho = -0.30$.

[Insert Figure A5 here]

Figure A6 shows sensitivity results for the analyses with Woodberry's (2012) Table 2, Model 3. The estimates for mass printing show some sensitivity. The ACME again becomes negative when ρ reaches 0.30. The mass education ACME becomes negative at $\rho = -0.30$.

[Insert Figure A6 here]

Finally, A7 shows sensitivity results for the model using our preferred specification. Not surprisingly, given that we reduced the number of covariates, it shows sensitivity more similar to the results in Figure A3. The mass printing and civil society estimates do not show much sensitivity, but the mass education estimates display some. Specifically, the mass education ACME becomes negative at $\rho = -0.15$.

[9] The graphs for the civil society mediator in Figures A3–A6 report Woodberry's (2012) Table 2, Model 3 specification. We could not estimate the models with $\rho \neq 0$ for that mediator using the main model due to singularity issues (see the replication materials).

[Insert Figure A7 here]

Overall, these results show sensitivity levels similar to other social science examples (see Imai et al. 2011; Imai and Yamamoto 2013). The estimates for the mass printing and civil society mediators, which are typically near zero, show the least amount of sensitivity. The estimates for mass education, in contrast, show a bit more. However, it is important to note that these analyses are not tests for the presence or absence of hidden confounders; they simply display sensitivity of the estimated mediation effects under varying strength of a hypothetical confounder. In this particular case, there is a wide variety of covariates included. Nonetheless, the results should be interpreted with appropriate caution.

6.2 Sensitivity to Treatment-Mediator Interaction

Our results that relax the assumption about association between mediators presented in Figure A2 require the assumption of no treatment-mediator interaction (Imai and Yamamoto 2013). Table A3 indicates that this assumption is reasonable. The table reports results from significance tests for treatment-mediator interactions, averaged across the 10 imputed datasets using the main model specification. Specifically, the first column reports the mean difference in ACME estimates under baseline differences in treatment status. The second column reports the minimum p-value associated with these differences and the third and fourth columns present the mean 95% confidence intervals. The differences are substantively small and statistically nonsignificant, lending credibility to the no interaction assumption.

[Insert Table A3 here]

Despite the results in Table A3, the no interaction assumption is technically applied to every observation (Imai and Yamamoto 2013, 157), so even if no interaction appears on average, the assumption could be violated for individual cases. Thus, we also report sensitivity analyses for that assumption in Figure A8. The graphs parameterize the hypothetical strength of treatment-mediator interaction with Imai and Yamamoto's (2013) $\tilde{R}^2$ term on the x-axes. That is, the x-axes plot the amount of variance that would be explained in the outcome if we could account for this heterogeneity. The solid lines on the y-axes indicate lower and upper bounds on the true ACME and shading indicates 95% confidence intervals. The dashed lines represent the estimated ACME under the no treatment-mediator interaction assumption.

[Insert Figure A8 here]

The top two graphs of Figure A8 indicate a high degree of sensitivity for the estimated mass printing ACMEs. The lower bounds reach zero with only a small amount of variance explained by a treatment-mediator interaction: $\tilde{R}^2 = 0.02$ when the alternative mediator is mass education and $\tilde{R}^2 = 0.01$ when it is civil society. The mass education estimates are more robust. When the alternative mediator is mass printing (civil society), the lower bound is zero at $\tilde{R}^2 = 0.10$ ($\tilde{R}^2 = 0.15$). For civil society, the values are $\tilde{R}^2 = 0.02$ (alternative: mass printing) and $\tilde{R}^2 = 0.07$ (alternative: mass education).

7 Mediator and Outcome Model Details

In this section we provide details on the models used to generate our mediation estimates.

7.1 Specifications

Table A4 reports the variables included in the models used to generate the estimates reported in the main text.

[Insert Table A4 here]

7.2 Results

Tables A5–A9 report output from the mediator and outcome models reported in Figure 1 of the main text and Figure A1. Consistent with Woodberry (2012), all coefficient estimates and standard errors come from robust regression (Street, Carroll, and Ruppert 1988).

[Insert Table A5 here]

[Insert Table A6 here]

[Insert Table A7 here]

[Insert Table A8 here]

[Insert Table A9 here]

8 Preregistered Robustness Checks

In this section we present results from several robustness checks described in our preanalysis plan.

8.1 Alternative Treatment Variables

To this point we have reported mediation results with Protestant missionaries per 10,000 population in 1923 as the treatment variable. In our preanalysis plan we recorded our intention to "repeat our main analyses with Woodberry's other primary variables of interest—percent evangelized by 1900 and years exposure to Protestant missionaries—as the treatment variable" (10). Figure A9 presents the results of those analyses. In each case we use the main model, but specify one of the other variables as the treatment.

[Insert Figure A9 here]

The top panel of Figure A9 presents results with percent evangelized by 1900. It indicates some similarities and some differences with the results shown in Figure 1 of the main text. The scale of this treatment variable is much different than the original treatment variable; it ranges from 0 to 100 with a mean of 41.32 and a standard deviation of 39.84. The original treatment measure ranges from 0 to 9.91 with a mean of 0.98 and a standard deviation of 1.67. As a result, the scale of the x-axis in this graph differs from Figure 1 in the main text. However, the magnitude of the total effect is comparable to our main results. An increase of one standard deviation in percent evangelized by 1900 corresponds with an expected increase of 7.6 on the democracy outcome measure. A standard deviation increase in the main treatment variable corresponds with a total effect of about 7.3.

Moving to the mediation effects, we see a relative pattern that is broadly similar to, but somewhat different from, our main results. The confidence intervals are sufficiently large such that none of the ACMEs are statistically significantly different from one another. However, the point estimates show important substantive variation. The education inequality variable again stands out as the strongest mediator; it produces a substantively large and statistically significant ACME estimate of 0.11, which is about 58% of the total effect. The civil society participation index is the second strongest, with an ACME of 0.04 (20% mediated). Newspaper circulation is the weakest of the three; the ACME estimate is 0.02, which corresponds to 9% mediated. Additionally, these latter two mediators' ACME estimates are not statistically significant at the 95% level.

The bottom panel of Figure A9, which gives results with years exposure to Protestant missionaries as the treatment, indicates a much different pattern from our main results. The total effect is a bit weaker and falls on the border of statistical significance at the 95% level. A one standard deviation increase in years exposure to Protestant missionaries (58.62) corresponds with an expected increase in democracy of 5.27. None of the ACME estimates reach statistical significance. The strongest of the three is the civil society participation index (0.029, 31% mediated). The ACME for newspaper circulation is negative and the estimate for education inequality is essentially zero.

In sum, the mediation analysis results display some sensitivity to the choice of treatment variable. However, as we described in our preanalysis plan, we chose Protestant missionaries per 10,000 population in 1923 because it is, in our view, the strongest measure of CP presence in a country. The percent evangelized measure includes converts to Catholicism in it (Woodberry 2012, 257) and years exposure measures the earliest time at which CPs were present in a country rather than levels of CPs in that country (Woodberry 2012, 263).[10] Accordingly, we place the most trust in the estimates using the Protestant missionaries per 10,000 population in 1923 variable.

8.2 Additional Covariates

In Figure A1 we show results after including two additional covariates: settler mortality (Acemoglu, Johnson, and Robinson 2001) and GDP per capita. Here we present results with each of those variables included on its own. The top panel of Figure A10 presents the mediation estimates

[10]Specifically, Woodberry (2012) measures this variable as $1960 - y_c$, where y_c is the first year in which CPs arrived in country c (263).

after adding settler mortality to the main model. The middle panel presents the same specification, but with GDP per capita included. Finally, the bottom panel reproduces the results from Figure A1 with both variables included.

[Insert Figure A10 here]

Substantive results largely remain the same after controlling for settler mortality. The total effect and ADE estimates are generally similar to those in Figure 1 of the main text, although the ADE estimated with education inequality weakens in magnitude and is no longer statistically significant at the 95% level. The ACME for education demonstrates a corresponding increase in magnitude—up to 42% of the total effect. The newspaper circulation and civil society participation mediators produce smaller ACMEs—14% and 7% of the total effect, respectively—and do not reach statistical significance.

In contrast to the results with settler mortality included, the results with only GDP added lead to a weakening of all three mediation effects. The middle panel of Figure A10 shows a reduction in magnitude for all three ACMEs. Additionally, none of them are statistically significant. Education still produces the strongest point estimate—about 24% of the total effect. Newspaper circulation mediates approximately 13% of the effect and the proportion mediated for civil society participation is 1%. However, we must interpret these results with caution because the GDP variable is measured posttreatment. Specifically, Woodberry (2012) computed it from World Bank data as the mean GDP per capita in each country beginning in 1960 (258).

8.3 Alternative Mediator Measures

In our own data collection efforts we gathered multiple indicators of each mediator with the intention of assessing robustness of our findings to the different measures (see section 3). We present results from some of these alternative measures in the main text and Figure A1. Here we report several more. Specifically, we present results from the main model specification for several additional combinations of mediator measures in Figure A11. As we show in Table A1, all of these indicators require multiple imputation to varying degrees. See section 9 for additional information.

Unless otherwise noted, the source of these indicators is V-Dem (Coppedge et al. 2018*a*).

[Insert Figure A11 here]

Beginning with the top row, the left graph shows results after replacing our education inequality and civil society participation measures with versions in which we compute the means over the period 1924–1949, which better matches the temporal sequencing of a mediator—after treatment (1923), but before measuring the outcome (1950). However, the main drawback to this strategy is that there are more missing data to impute. With our original versions (averaging from 1924–1994), 29% and 12% of the education and civil society mediators are missing, respectively. Those numbers increase to 54% and 19% with the 1924–1949 measures. The change is not consequential for the civil society participation index—its ACME is still close to zero and nonsignificant. Education inequality's ACME increases slightly in magnitude from the main text results (37%). However, the added uncertainty that stems from more missing data renders that estimate not significant at the 95% level.

The middle panel uses our book titles measure of mass printing, averaged over the period 1924–1949. To maintain consistency within the graph, we also use the versions of education inequality and the participation index measured over that period. The estimated ACME for book titles is similar to the 1924–1994 version reported in the main text: substantively fairly small and statistically not significant. The right panel reports results using the literacy measure averaged over 1924–1949 for education. The ACME estimate is slightly larger than the estimate using 1924–1994 data (main text and Figure A1), but the estimate is statistically nonsignificant.

Next, we replace the education mediator with Woodberry's (2012) education measure: secondary school enrollment, averaged over 1960–1985 (bottom left panel). Importantly, this variable reflects another conceptual change in measurement. The inequality measure captures Woodberry's (2012) theoretical discussion about whether access to education is reserved for elites or widely available (e.g., 246). The secondary school measure focuses more on average levels of education. This choice is fairly consequential for the results. The education ACME drops considerably in magnitude compared to the inequality measure (18% of the total effect) and is no longer statistically significant.

The bottom middle and left panels show results with education measured as mean years of education (averaged 1924–1994 and 1924–1949). This measure is conceptually similar to the secondary school enrollment measure, although we are able to average it over longer periods of time. The estimated ACME in the middle graph (1924–1994) is similar to what we report in the main text with respect to substantive magnitude and statistical significance. However, the estimate drops in magnitude and loses significance when the variable is measured over 1924–1949. Overall, these results reinforce the point that substantive results are somewhat contingent on the mediator measure chosen.

8.4 Subsetting by Region

In our preanalysis plan we declared our intent to "repeat our main analyses in several regions of the world discussed by Woodberry" (11). Unfortunately, we were not successful in doing so due to relatively small sample sizes in the subsets. Subsetting to the various regions yielded samples of 40–50 countries, which frequently produced singularities and other estimation issues. Even after removing problematic variables from the model specifications, results were nonsensical and/or noninformative. Most frequently, the bootstrapping procedure produced extremely large confidence intervals. We also experienced this problem when we tried to include region $\times$ treatment interaction terms in the models as part of our preferred specification. For the sake of completeness, we include the code for these analyses in the replication materials. However, our assessment is that estimating Woodberry's model on regional subsets is asking too much of these data.

8.5 Subsetting by Colonizer

Our preanalysis plan also mentions subsetting by colonizing country. We had more success with these subsets because they are slightly larger. Specifically, we repeated our main analyses separately for British colonies (N = 50), countries colonized by Protestant countries (N = 57), countries colonized by countries that are not predominantly Protestant (N = 75), and countries colonized by Catholic countries (N = 58).[11] Figure A12 reports results for these subsets using the main model specification.[12]

[Insert Figure A12 here]

One obvious feature of these graphs is the increase in confidence interval size due to the smaller subsets. None of the reported estimates are statistically significantly different from each other, so we must interpret them with some caution. Nonetheless, it is still instructive to consider the magnitudes and signs of the point estimates. Beginning with British colonies in the top left panel, note that the total effect is positive, but reduced to about 40% of its value in our main results. The ACMEs suggest that mass printing and mass education are roughly equal in their mediation effects (both about 24% of the total effect). The civil society ACME is very close to zero. The result changes when we consider Protestant colonies (top right panel). There the ACME for education inequality emerges as the largest (52% mediated) and newspaper circulation declines (11%). Additionally, the civil society participation index estimate moves away from zero, mediating 28% of the total effect.

The picture changes more drastically in the bottom two panels of Figure A12. Moving outside of Protestant colonies to countries that were colonized by any non-Protestant country (bottom left) or Catholic countries specifically (bottom right) renders the total effect negative and nonsignificant. The relative ordering of the mediation effects remains the same, but these results show that the relationship between CPs and democracy itself is in question in these subsets of countries.

8.6 Omitting Multiple Imputation

Missing data in our mediator variables necessitate that we use multiple imputation to produce complete data and avoid the potential for bias from listwise deletion (e.g., Blackwell, Honaker, and King 2017). However, we also repeated our main analyses using listwise deletion to assess dependence on imputation. The results, which appear in Figure A13, are generally similar to our

[11] This list represents the set of feasible colonizer subsets. It is slightly different from the list we planned to use in our preanalysis plan (11).

[12] These subsets required us to remove some variables, such as the indicator for British colonies, from the model specification. See the replication materials for more details.

results with imputed data, but with some key differences that support our decision to impute missing data. The three total effect estimates display more heterogeneity due to the fact that they are estimated on different samples depending on which data are observed for each mediator. None of the three samples recover Woodberry's (2012) original estimate of 4.43 (as our multiple imputation analysis does), although the mass education sample comes close (4.25). However, there is enough uncertainty in the estimates such that two of the three samples (mass printing and mass education) produce a statistically nonsignificant total effect. Thus, without multiple imputation we are not providing a comparable replication and extension of the original results.

The relative ordering of the ACMEs remains the same as in our main results, although the magnitudes of each are larger with listwise deletion. According to these results, the three mediators account for 23% (newspaper circulation), 59% (education inequality), and 7% (civil society participation) of the estimated total effects. As in our main text results, only the mass education estimate is statistically significantly different from zero. Overall, these results tell a fairly similar story to what we report using imputation. But we place less trust in these estimates (and more in our results using imputed data) given the problems that arise with listwise deletion (Blackwell, Honaker, and King 2017).

[Insert Figure A13 here]

9 Multiple Imputation Diagnostics

We used multiple imputation with the Amelia II software available in R (Honaker, King, and Blackwell 2011) to fill in missing data values. Amelia begins with the assumption that the complete data are distributed multivariate normal, then employs the expectation-maximization (EM) algorithm to impute the missing values. It repeats this process m times, then analysis continues as usual with those m complete datasets. In our preanalysis plan (8) we indicated the intention to use the default value of $m = 5$. However, the imputed datasets evidenced fairly substantial heterogeneity, especially when imputing variables with larger proportions of missingness. In response to this issue, we increased m to 10 for all of the results presented in this research. We set bounds on each imputed variable to eliminate impossible values. Additionally, we set weakly informative normal priors on imputed variables. For each one we set the prior mean to the mean of the observed values and the prior standard deviation to five times the standard deviation of the observed values.[13]

An adjustment to measures of uncertainty is necessary when imputing data (see Blackwell, Honaker, and King 2017, 309). We accounted for the additional variance that stems from imputation by following the steps discussed in Blackwell, Honaker, and King (2017, 309, see also Imai, King, and Lau 2008). First, we simulated our quantities of interest (ACMEs, ADEs, and total effects) from the mediation models 100 times from each of the $m = 10$ datasets via bootstrapping. Then we combined the $100 \times 10 = 1,000$ replicates as if they came from the same model. The results presented in this research come from summarizing the distributions of each quantity of interest over the 1,000 bootstrap replicates.

Amelia II provides two key diagnostic tools for evaluating the quality of imputations: overimputation and density plots. The former conducts imputation of the observed data, then compares the imputed to the actual values of those data. The latter involves graphing the distributions of observed and imputed values of each variable. We report these diagnostics for the variables we imputed below, beginning with the mediators.

9.1 Main Mediator Measures

Figure A14 presents overimputation results for our three main mediator measures. In each graph, the observed values of the non-missing data points are plotted on the x-axes and imputed values (averaged over the 10 datasets) of those data are plotted on the y-axes. The vertical line segments indicate 95% confidence intervals for the imputations and the solid line serves as a reference point for "perfect" imputation. In an ideal scenario the points would fall along the reference line. More realistically, favorable evidence for the imputation procedure would exist if (approximately) 95% of the confidence intervals include the reference line. The colors classify each point based on this criterion: blue indicates points for which the confidence interval includes the reference line

[13] The use of these priors improved imputation quality (as measured by overimputation—see below), but does not affect substantive conclusions from the mediation results.

and red indicates points that do not.

[Insert Figure A14 here]

The graphs in Figure A14 show good, though not perfect, coverage of the reference line. The clouds of points trend upward with the line, and most of the points are blue. The actual coverage rates are slightly less than, but close to the target: 92% (newspaper circulation), 90% (education inequality), and 92% (civil society participation). Thus, the imputation results fall short of ideal, but are nonetheless reasonable. Additionally, it is important to keep in mind that the use of multiple imputation is not wildly consequential for our substantive conclusions (see Figure A13).

Figure A15 presents density plots of the observed (blue) and imputed (red) values (averaged across the 10 datasets) of each mediator variable. These graphs indicate considerable overlap between the two groups. Thus, there is evidence that the imputation procedure produced reasonable values for the missing data.

[Insert Figure A15 here]

9.2 Alternative Mediator Measures

Table A10 reports the percent missing and overimputation coverage rates for each of the alternative mediator measures. These rates are computed using 95% confidence intervals, so as in the graphs discussed above, the favorability of the imputation procedure increases as the rate gets closer to 95%. Overall, these rates are close to 95%, suggesting that multiple imputation worked well for these data.

[Insert Table A10 here]

9.3 Additional Covariates

Figures A16 and A17 present overimputation results and density plots for the two covariates we imputed: settler mortality and GDP. Both show similar patterns to the graphs of the mediator measures. Figure A16 indicates that most of the overimputations follow the reference line (the actual coverage rates are 97% for settler mortality and 89% for GDP). Figure A17 shows considerable overlap between the observed and average imputed values.

[Insert Figure A16 here]

[Insert Figure A17 here]

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development." *American Economic Review* 91(5): 1369–1401.

Baron, Reuben M., and David A. Kenny. 1986. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51(6): 1173–1182.

Barro, Robert J., and Jong-Wha Lee. 1994. "Data Set for a Panel of 138 Countries." http://www.nber.org/pub/barro.lee/.

Blackwell, Matt, James Honaker, and Gary King. 2017. "A Unified Approach to Measurement Error and Missing Data: Overview and Applications." *Sociological Methods & Research* 46(3): 303–341.

Bollen, Kenneth A. 2009. "Liberal Democracy Series I, 1972–1988." *Electoral Studies* 28(3): 368–374.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jerey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig, and Daniel Ziblatt. 2018*a*. "V-Dem Codebook v8." Varieties of Democracy (V-Dem) Project. https://www.v-dem.net/.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jerey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig, and Daniel Ziblatt. 2018*b*. "V-Dem [Country-Year/Country-Date] Dataset v8." Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemcy18.

Fink-Jensen, Jonathan. 2015. "Book Titles per Capita." IISH Dataverse. http://hdl.handle.net/10622/AOQMAZ.

Haavelmo, Trygve. 1943. "The Statistical Implications of a System of Simultaneous Equations." *Econometrica* 11(1): 1–12.

Honaker, James, Gary King, and Matthew Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45(7): 1–47.

Imai, Kosuke, and Teppei Yamamoto. 2013. "Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments." *Political Analysis* 21(2): 141–171.

Imai, Kosuke, Gary King, and Olivia Lau. 2008. "Toward a Common Framework for Statistical Analysis and Development." *Journal of Computational Graphics and Statistics* 17(4): 1–22.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. "Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25(1): 51–71.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational

Studies." *American Political Science Review* 105(4): 765–789.

Keele, Luke, Dustin Tingley, and Teppei Yamamoto. 2015. "Identifying Mechanisms Behind Policy Interventions via Causal Mediation Analysis." *Journal of Policy Analysis and Management* 34(4): 937–963.

Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. "How Conditioning on Post-treatment Variables Can Ruin Your Experiment and What to Do about It." Forthcoming, *American Journal of Political Science*.

Paxton, Pamela. 2002. "Social Capital and Democracy: An Interdependent Relationship." *American Sociological Review* 67(2): 254–277.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi ting Wang, Joshua Krusell, and Farhad Miri. 2018. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21, 3d edition.

Street, James O., Raymond J. Carroll, and David Ruppert. 1988. "A Note on Computing Robust Regression Estimates Via Iteratively Reweighted Least Squares." *The American Statistician* 42(2): 152–154.

Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele, and Kosuke Imai. 2014. "mediation: R Package for Causal Mediation Analysis." *Journal of Statistical Software* 59(5): 1–38.

Vanhanen, Tatu. 2003. *Democratization: A Comparative Analysis of 170 Countries*. New York: Routledge.

Woodberry, Robert D. 2012. "The Missionary Roots of Liberal Democracy." *American Political Science Review* 106(2): 244–274.

Table A1: Missingness in the Measures of Causal Mechanisms

Mediator	Variable	Source	Cases Missing	% Missing
Mass Printing	Mean daily newspaper circulation (average of 1975, 1980, 1985, 1990)	Woodberry (2012), UN Data	24	16.9
	Mean book titles per capita (average)	Fink-Jensen (2015)	129/112	90.8/78.9*
Mass Education	% Literate (average)	Vanhanen (2003); Coppedge et al. (2018*a*)	96/32*	67.6/22.5*
	Mean years education (average)	Coppedge et al. (2018*a*)	77/42*	54.2/29.6*
	Secondary school enrollment (1960–1985)	Woodberry (2012)	57	40.1
	Education inequality (average)	Coppedge et al. (2018*a*)	76/42*	53.5/29.6*
Civil Society	Civil society participation index (average)	Coppedge et al. (2018*a*)	27/18*	19.0/12.7*

Note: Cell entries summarize missing data in the mediator measures. The percentage missing for each variable is computed based on the 142 countries in the estimation sample of Woodberry's (2012) Table 3, Model 4 (262). * Pairs of entries denoted ·/· indicate the two measurement strategies discussed in the text: computing the averages from 1924–1949 (listed first) or 1924–1994 (listed second).

Table A2: Regression Models Testing Associations Between Mediators

Outcome	Newspaper Circulation	Education Inequality	Participation Index
Newspaper Circulation		-0.12^*	0.001^*
		(0.02)	(0.000)
Education Inequality	-2.03^*		-0.001
	(0.44)		(0.001)
Participation Index	116.90^*	-5.52	
	(39.12)	(10.83)	
Treatment	-3.21	-2.89^*	-0.001
	(6.87)	(1.27)	(0.01)
N	142	142	142
Adjusted R^2	0.59	0.78	0.37

Note: Cell entries report ordinary least squares (OLS) coefficient estimates with standard errors in parentheses from models in which one mediator is regressed on the alternative mediators, treatment, and the covariates (using the main model specification). Results reflect combination of the 10 imputed datasets. * $p < 0.05$.

Table A3: Significance Tests for Treatment-Mediator Interactions

Mediator	Mean Difference	Minimum p-value	95% Confidence Interval	
			Lower	Upper
Mass Printing	−0.048	0.360	−0.389	0.175
Mass Education	0.146	0.200	−0.256	0.687
Civil Society	0.003	0.400	−0.422	0.316

Note: Cell entries report results of significance tests for treatment-mediator interaction for each mediator, averaged across the imputed datasets. The first column reports the mean difference in ACME estimates under baseline differences in treatment status. The second column reports the minimum p-value associated with these differences and the third and fourth columns present the mean 95% confidence intervals.

Table A4: Model Specifications Reported in the Main Text

Main Model: Table 3, Model 4 (Woodberry 2012, 262)	Additional Covariates: Settler Mortality and GDP	Alternative Printing and Education Measures	Table 2, Model 3 (Woodberry 2012, 260)	Authors' Preferred Specification
• British Colony • Other Religious Liberty Colony • Dutch Colony • Never Colonized Significantly • Latitude • Island Nation • Landlocked Nation • Percent European in 1980 • Percent Muslim in 1970 • Major Oil Producer • Literate Culture Before Missionary Contact • Years Exposure to Protestant Missions • Protestant Missionaries per 10,000 population in 1923 (Treatment) • Percent Evangelized by 1900 • Years Exposure to Catholic Missions • Foreign Catholic Priests per 10,000 population in 1923 • Year of 1st Democracy Data • Post-1976 Democracy Data Only • Date 1st Sighted by Europeans after 1444 • Gap between Sighted and 1st Missionaries • Mission Gap × Literacy • Mission Gap × Latitude • Gap between Sighted and Colonized • Colonial Gap × Literacy • Colonial Gap × Latitude • Number of Times Territory Switched Colonizers • Protestant Colonizer Took Colony from Catholics • *Average Newspaper Circulation* • *Education Inequality* • *Civil Society Participation Index*	• British Colony • Other Religious Liberty Colony • Dutch Colony • Never Colonized Significantly • Latitude • Island Nation • Landlocked Nation • Percent European in 1980 • Percent Muslim in 1970 • Major Oil Producer • Literate Culture Before Missionary Contact • Years Exposure to Protestant Missions • Protestant Missionaries per 10,000 population in 1923 (Treatment) • Percent Evangelized by 1900 • Years Exposure to Catholic Missions • Foreign Catholic Priests per 10,000 population in 1923 • Year of 1st Democracy Data • Post-1976 Democracy Data Only • Date 1st Sighted by Europeans after 1444 • Gap between Sighted and 1st Missionaries • Mission Gap × Literacy • Mission Gap × Latitude • Gap between Sighted and Colonized • Colonial Gap × Literacy • Colonial Gap × Latitude • Number of Times Territory Switched Colonizers • Protestant Colonizer Took Colony from Catholics • Settler Mortality • ln(GDP per Capita) • *Average Newspaper Circulation* • *Education Inequality* • *Civil Society Participation Index*	• British Colony • Other Religious Liberty Colony • Dutch Colony • Never Colonized Significantly • Latitude • Island Nation • Landlocked Nation • Percent European in 1980 • Percent Muslim in 1970 • Major Oil Producer • Literate Culture Before Missionary Contact • Years Exposure to Protestant Missions • Protestant Missionaries per 10,000 population in 1923 (Treatment) • Percent Evangelized by 1900 • Years Exposure to Catholic Missions • Foreign Catholic Priests per 10,000 population in 1923 • Year of 1st Democracy Data • Post-1976 Democracy Data Only • Date 1st Sighted by Europeans after 1444 • Gap between Sighted and 1st Missionaries • Mission Gap × Literacy • Mission Gap × Latitude • Gap between Sighted and Colonized • Colonial Gap × Literacy • Colonial Gap × Latitude • Number of Times Territory Switched Colonizers • Protestant Colonizer Took Colony from Catholics • *Book Titles per Capita* • *Percent Literate* • *Civil Society Participation Index*	• British Colony • Other Religious Liberty Colony • Dutch Colony • Never Colonized Significantly • Latitude • Island Nation • Landlocked Nation • Percent European in 1980 • Percent Muslim in 1970 • Major Oil Producer • Literate Culture Before Missionary Contact • Years Exposure to Protestant Missions • Protestant Missionaries per 10,000 population in 1923 (Treatment) • Percent Evangelized by 1900 • Years Exposure to Catholic Missions • Foreign Catholic Priests per 10,000 population in 1923 • Year of 1st Democracy Data • Post-1976 Democracy Data Only • *Average Newspaper Circulation* • *Education Inequality* • *Civil Society Participation Index*	• British Colony • Other Religious Liberty Colony • Dutch Colony • Never Colonized Significantly • Latitude • Island Nation • Landlocked Nation • Major Oil Producer • Literate Culture Before Missionary Contact • Protestant Missionaries per 10,000 population in 1923 (Treatment) • Years Exposure to Catholic Missions • Foreign Catholic Priests per 10,000 population in 1923 • Year of 1st Democracy Data • Post-1976 Democracy Data Only • Protestant Colonizer Took Colony from Catholics • Settler Mortality • Region Indicator Variables (Sub-Saharan Africa, Asia, Latin America and the Caribbean, the Middle East/North Africa, and Oceania) • *Average Newspaper Circulation* • *Education Inequality* • *Civil Society Participation Index*

Note: Cell entries report the covariates included in the models used to generate the estimates reported in the main text. Mediators are listed in italics. The outcome variable is the average democracy score during 1950–1994 from the Cross-national Indicators of Liberal Democracy series (see Paxton 2002; Bollen 2009).The main model refers to Table 3, Model 4 in Woodberry (2012, 262). Region indicators × Treatment interaction terms were intended to be included in the preferred specification, but were ultimately dropped due to estimation problems (see section 8.4).

Table A5: Mediator and Outcome Model Output with the Main Model

Variable	Mass Printing		Mass Education		Civil Society	
	Mediator	Outcome	Mediator	Outcome	Mediator	Outcome
Newspaper Circulation		0.12				
		(0.03)				
Education Inequality				−0.43		
				(0.13)		
Participation Index						48.27
						(11.98)
Treatment	4.63	3.89	−3.34	2.99	0.00	4.26
	(6.13)	(1.39)	(1.28)	(1.38)	(0.01)	(1.32)
N	142	142	142	142	142	142
Adjusted R^2	0.40	0.64	0.73	0.64	0.31	0.65

Note: Cell entries report robust regression coefficient estimates with standard errors in parentheses for the models reported in Figure 1 of the main text. Results reflect combination of the 10 imputed datasets.

Table A6: Mediator and Outcome Model Output with the Additional Covariates

Variable	Mass Printing		Mass Education		Civil Society	
	Mediator	Outcome	Mediator	Outcome	Mediator	Outcome
Newspaper Circulation		0.11				
		(0.03)				
Education Inequality				−0.42		
				(0.13)		
Participation Index						52.13
						(12.88)
Treatment	7.16	3.58	−3.25	3.05	0.01	4.08
	(7.57)	(1.54)	(1.57)	(1.47)	(0.01)	(1.43)
N	142	142	142	142	142	142
Adjusted R^2	0.38	0.64	0.72	0.64	0.38	0.66

Note: Cell entries report robust regression coefficient estimates with standard errors in parentheses for the models reported in the top left panel of Figure 2 of the main text. Results reflect combination of the 10 imputed datasets.

Table A7: Mediator and Outcome Model Output with the Alternative Mediator Measures

Variable	Mass Printing		Mass Education		Civil Society	
	Mediator	Outcome	Mediator	Outcome	Mediator	Outcome
Book Titles per Capita		0.01				
		(0.01)				
Percent Literate				0.34		
				(0.14)		
Participation Index						47.90
						(12.32)
Treatment	21.19	3.91	1.55	3.90	0.01	4.11
	(12.70)	(12.20)	(1.47)	(1.33)	(0.01)	(1.27)
N	142	142	142	142	142	142
Adjusted R^2	0.82	0.61	0.70	0.63	0.30	0.65

Note: Cell entries report robust regression coefficient estimates with standard errors in parentheses for the models reported in the top right panel of Figure 2 of the main text. Results reflect combination of the 10 imputed datasets. The results for Civil Society are slightly different from those reported in Table A5 because the data reflect a new iteration of imputation with the alternative mediator measures for mass printing and mass education.

Table A8: Mediator and Outcome Model Output with Woodberry's (2012) Table 2, Model 3 Specification

Variable	Mass Printing		Mass Education		Civil Society	
	Mediator	Outcome	Mediator	Outcome	Mediator	Outcome
Newspaper Circulation		0.09				
		(0.03)				
Education Inequality				−0.36		
				(0.14)		
Participation Index						52.18
						(12.49)
Treatment	8.69	2.95	−2.46	2.90	0.01	3.41
	(6.79)	(1.41)	(1.23)	(1.32)	(0.01)	(1.27)
N	142	142	142	142	142	142
Adjusted R^2	0.36	0.58	0.68	0.59	0.33	0.61

Note: Cell entries report robust regression coefficient estimates with standard errors in parentheses for the models reported in the bottom left panel of Figure 2 of the main text. Results reflect combination of the 10 imputed datasets.

Table A9: Mediator and Outcome Model Output with the Authors' Preferred Specification

Variable	Mass Printing		Mass Education		Civil Society	
	Mediator	Outcome	Mediator	Outcome	Mediator	Outcome
Newspaper Circulation		0.03				
		(0.03)				
Education Inequality				−0.13		
				(0.10)		
Participation Index						29.86
						(10.78)
Treatment	4.85	4.14	−3.68	3.80	0.00	4.17
	(6.23)	(1.12)	(1.59)	(1.15)	(0.01)	(1.10)
N	142	142	142	142	142	142
Adjusted R^2	0.35	0.71	0.67	0.71	0.37	0.72

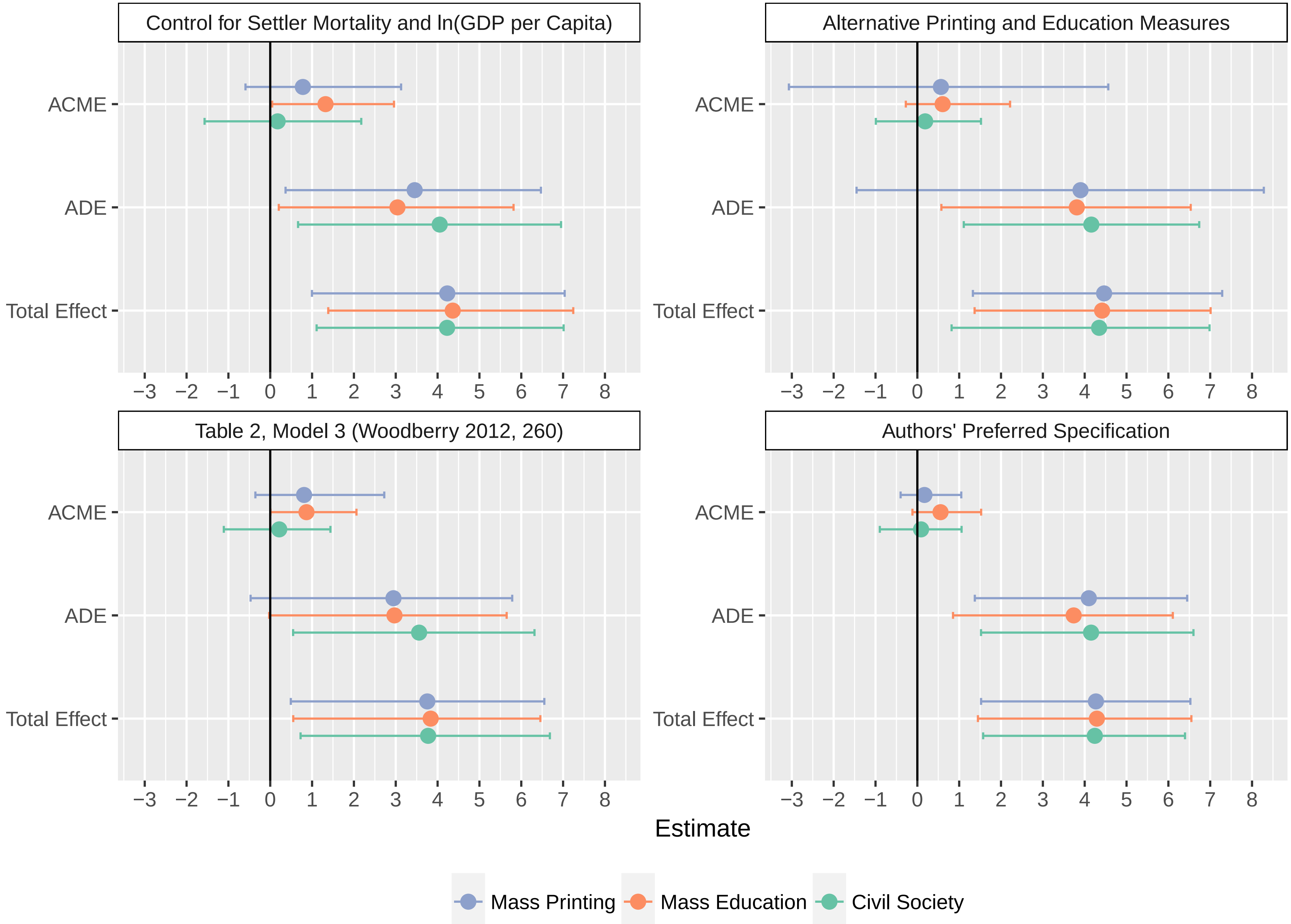
Note: Cell entries report robust regression coefficient estimates with standard errors in parentheses for the models reported in the bottom right panel of Figure 2 of the main text. Results reflect combination of the 10 imputed datasets.

Table A10: Overimputation Coverage Rates for the Alternative Mediator Measures

Mediator	Variable	Percent Missing	Coverage
Mass Printing	Book Titles per Capita (1924–1994)	79%	97%
	Book Titles per Capita (1924–1949)	91%	100%
Mass Education	Education Inequality (1924–1949)	54%	92%
	Percent Literate (1924–1994)	23%	92%
	Percent Literate (1924–1949)	68%	91%
	Secondary School Enrollment (1960–1985)	40%	95%
	Mean Years of Education (1924–1994)	30%	91%
	Mean Years of Education (1924–1949)	54%	92%
Civil Society	Participation Index (1924–1949)	19%	89%

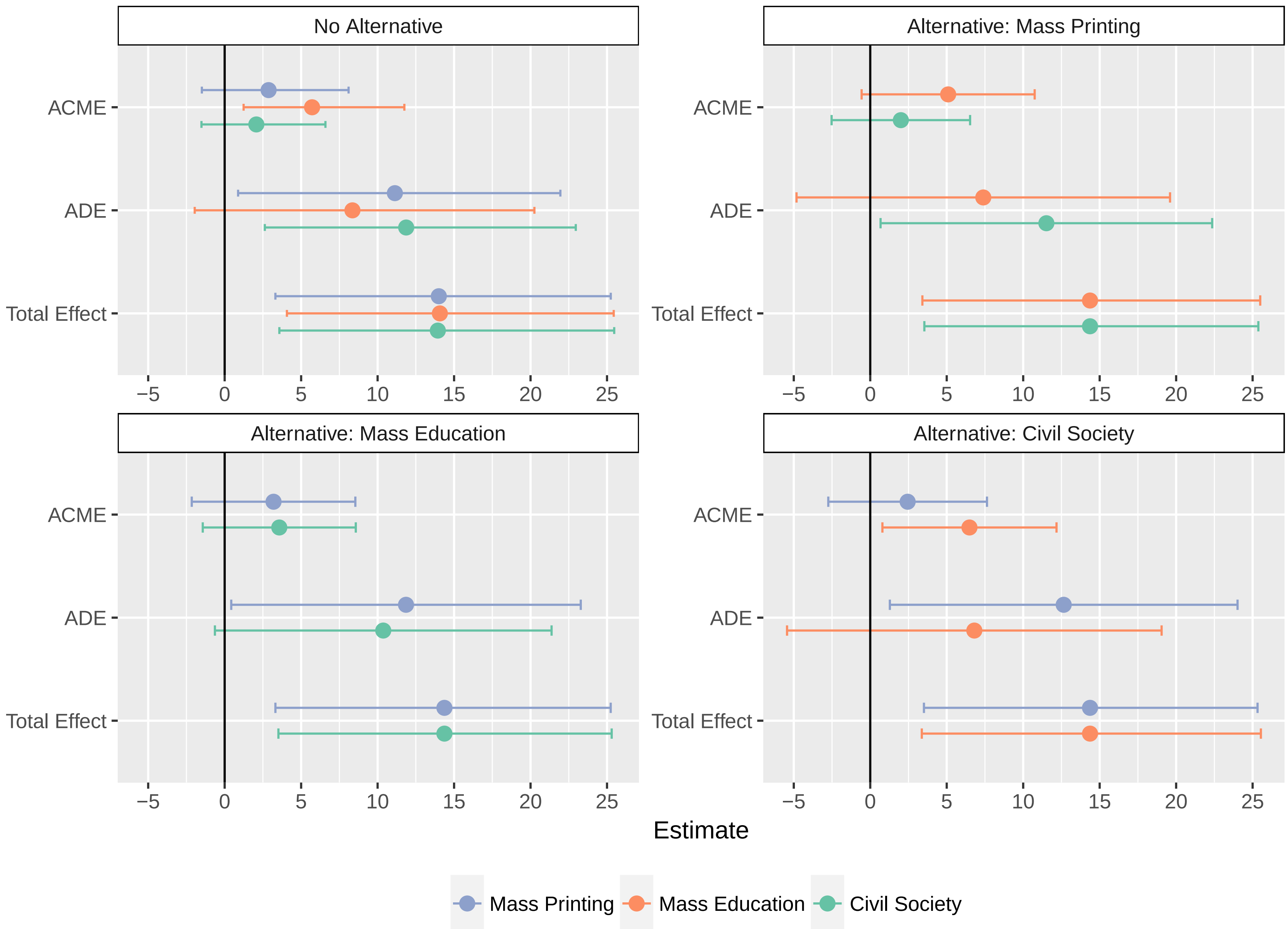
Note: Cell entries report percent missing and overimputation coverage rates (using 95% confidence intervals) for each of the alternative mediator measures.

Figure A1: Estimated Mediation Effects with Alternative Model Specifications



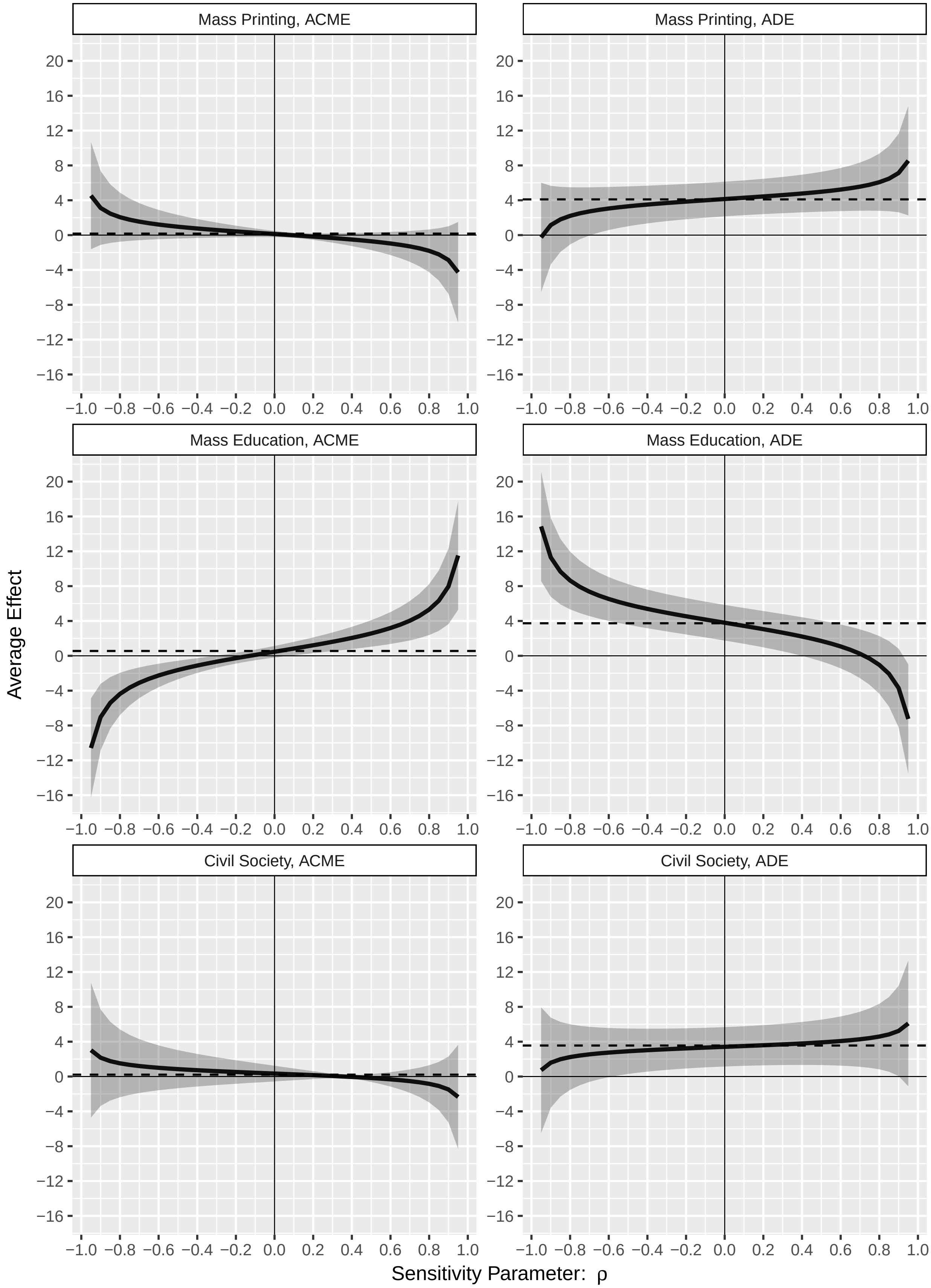
Note: The graphs present the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals. The top left panel presents results after adding settler mortality (Acemoglu, Johnson, and Robinson 2001) and the natural log of GDP per capita to the main model. The top right panel presents results with book titles per capita as the measure of mass printing and percent literate as the measure of mass education. Results in the bottom left panel come from another specification to which Woodberry (2012) frequently refers (Table 2, Model 3, 260), using the original mediator measures. Finally, the bottom right panel presents results from our preferred specification (using the original mediators), which addresses posttreatment bias, streamlines the definition of treatment, and conserves degrees of freedom.

Figure A2: Estimated Mediation Effects Controlling for Association Between Mediators



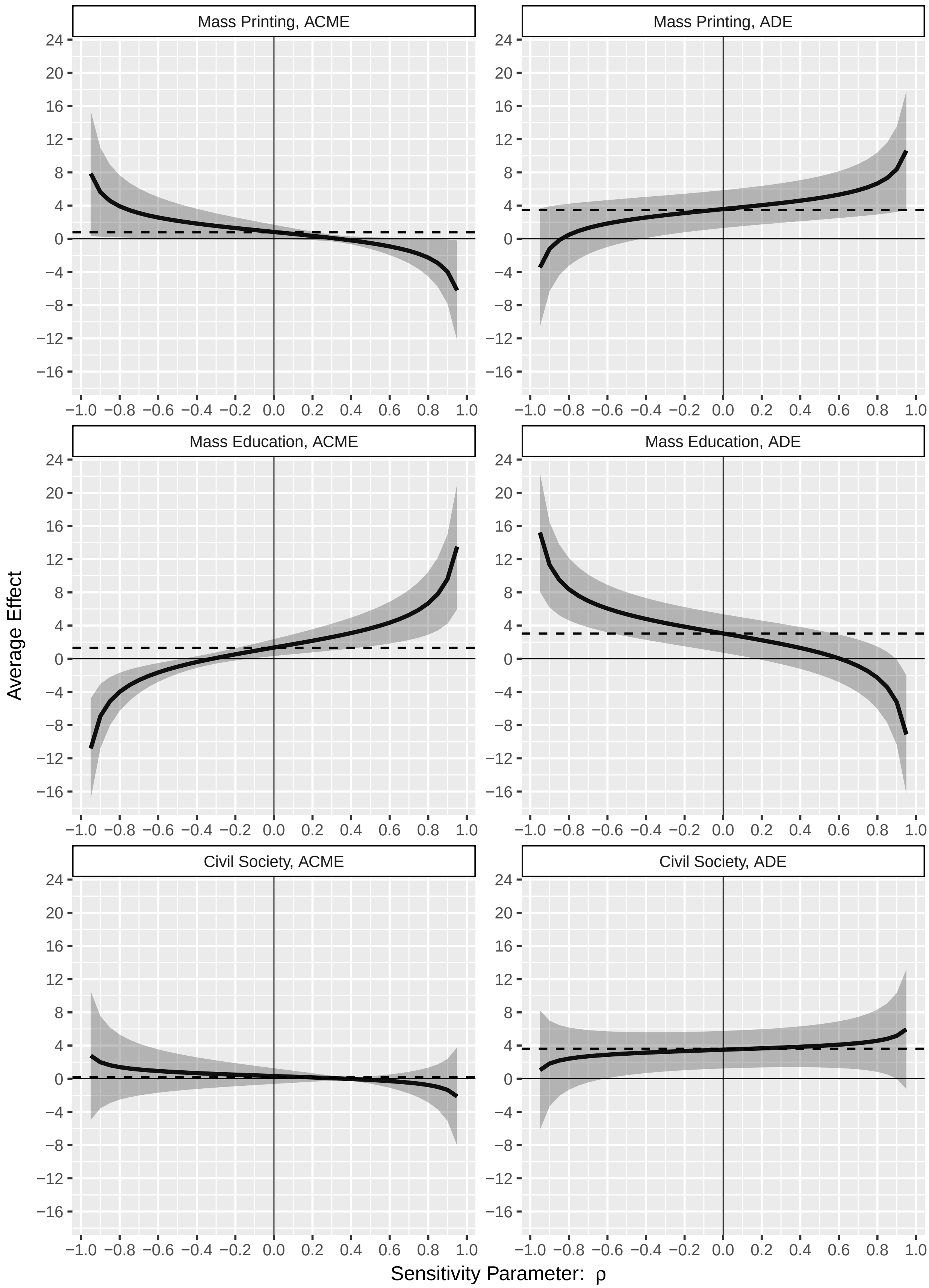
Note: The graphs present the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals. The top left panel is analogous to Figure 1 in the main text, but with a binary version of the treatment variable. The other three panels report estimates that control for association with the alternative mediator listed at the top of each graph. For example, the first estimate in the top right panel is the ACME for mass education after controlling for the possible confounding influence of mass printing.

Figure A3: Sensitivity Analysis of Confounding by Pretreatment Covariates in the Main Model



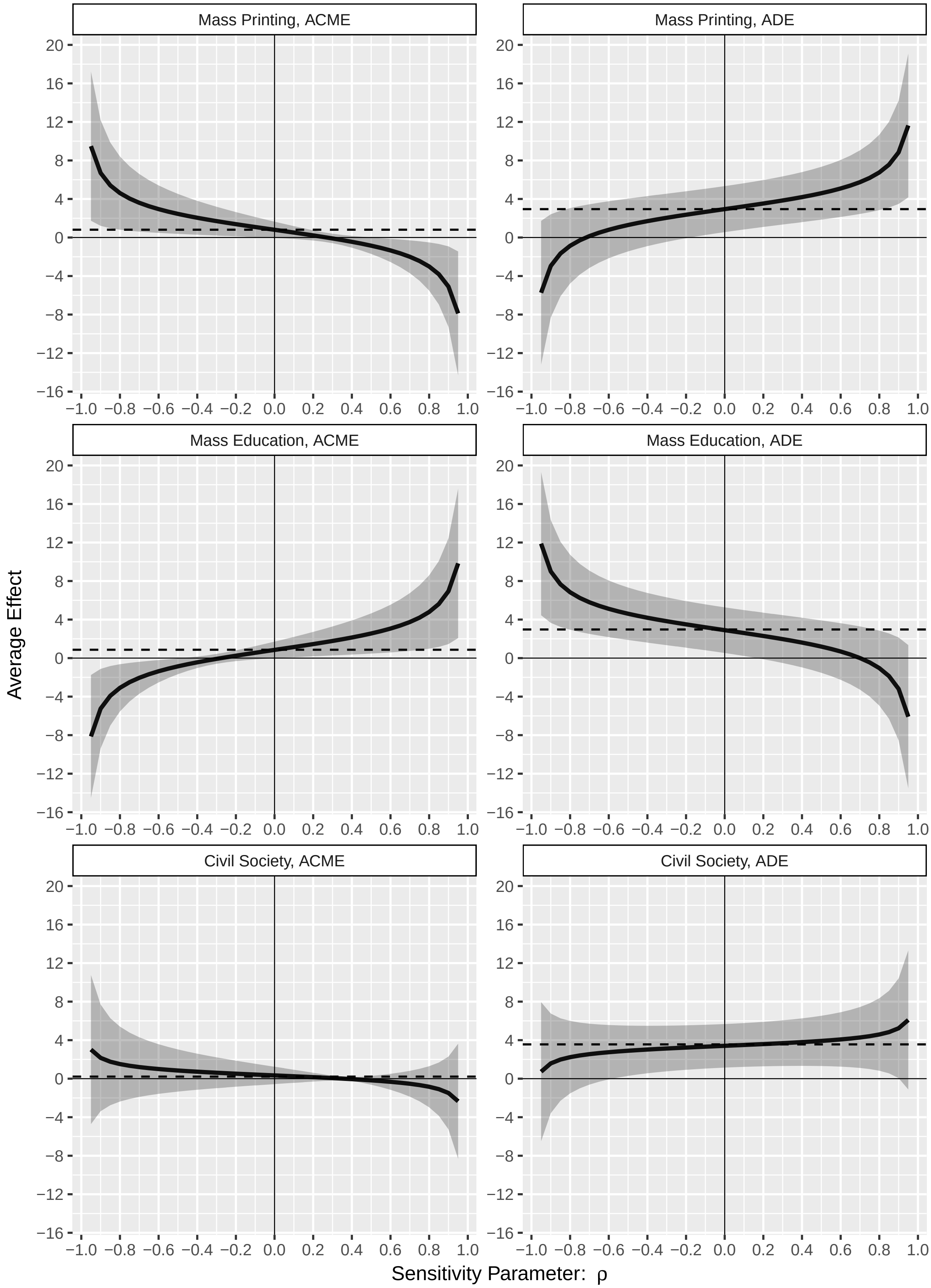
Note: The graphs present, for each mechanism, the sensitivity of the estimated ACMEs and ADEs to hidden confounding from an omitted pretreatment covariate. The dashed lines indicate the estimates assuming no hidden confounder.

Figure A4: Sensitivity Analysis of Confounding by Pretreatment Covariates After Controlling for Settler Mortality and GDP



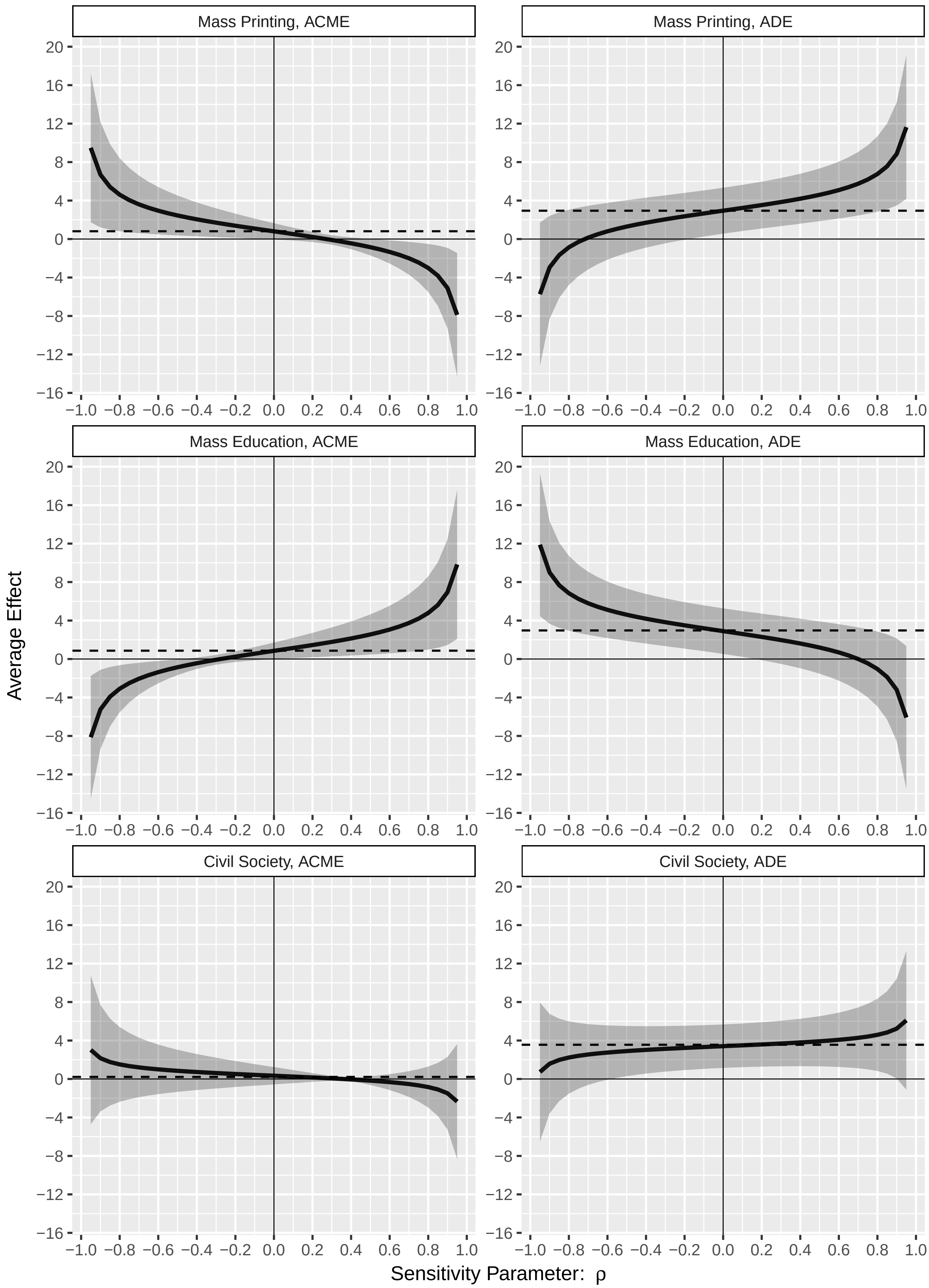
Note: The graphs present, for each mechanism, the sensitivity of the estimated ACMEs and ADEs to hidden confounding from an omitted pretreatment covariate. The dashed lines indicate the estimates assuming no hidden confounder.

Figure A5: Sensitivity Analysis of Confounding by Pretreatment Covariates with Alternative Printing and Education Mediator Measures



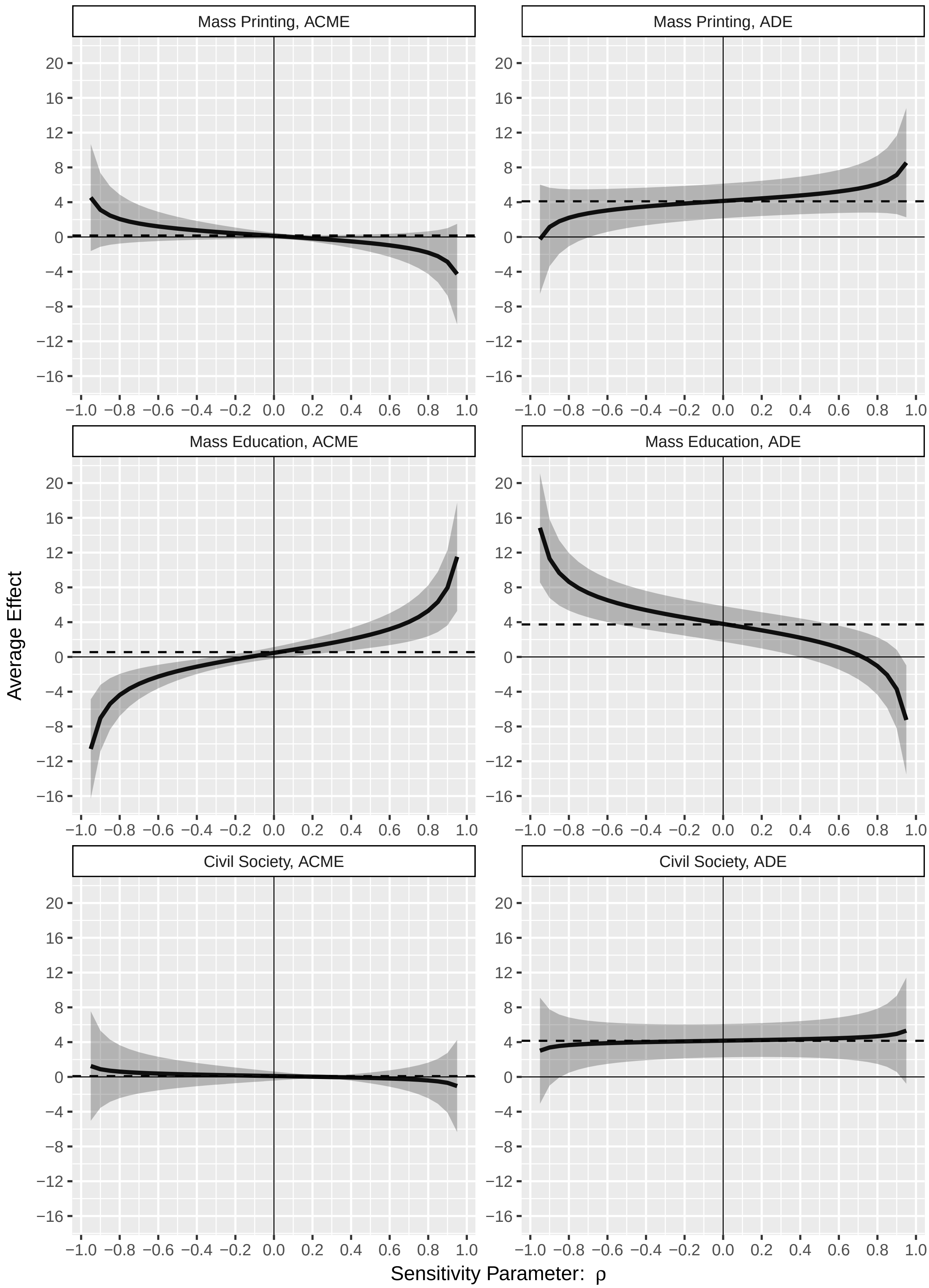
Note: The graphs present, for each mechanism, the sensitivity of the estimated ACMEs and ADEs to hidden confounding from an omitted pretreatment covariate. The dashed lines indicate the estimates assuming no hidden confounder.

Figure A6: Sensitivity Analysis of Confounding by Pretreatment Covariates with Woodberry's (2012) Table 2, Model 3 Specification



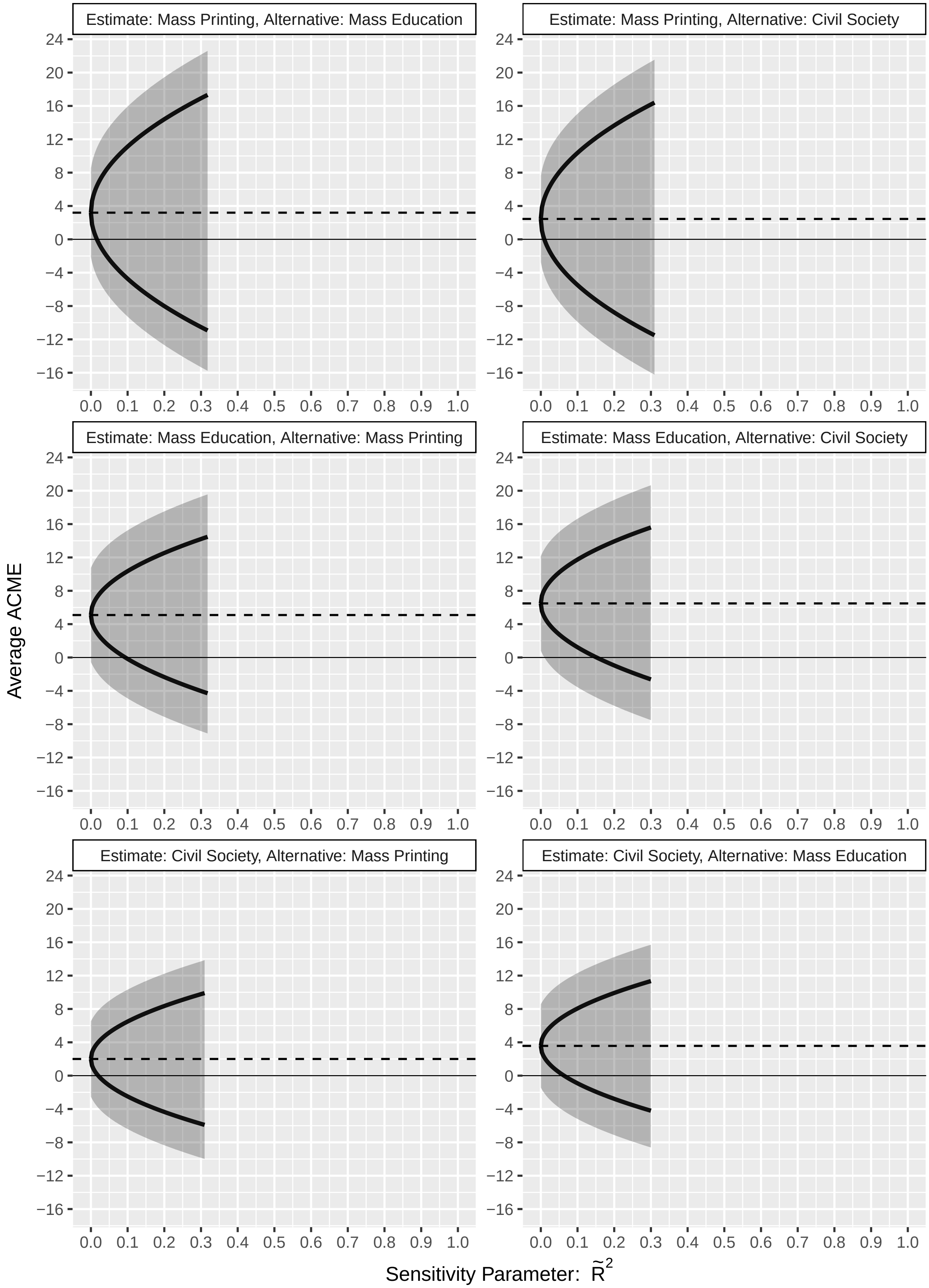
Note: The graphs present, for each mechanism, the sensitivity of the estimated ACMEs and ADEs to hidden confounding from an omitted pretreatment covariate. The dashed lines indicate the estimates assuming no hidden confounder.

Figure A7: Sensitivity Analysis of Confounding by Pretreatment Covariates with Authors' Preferred Specification



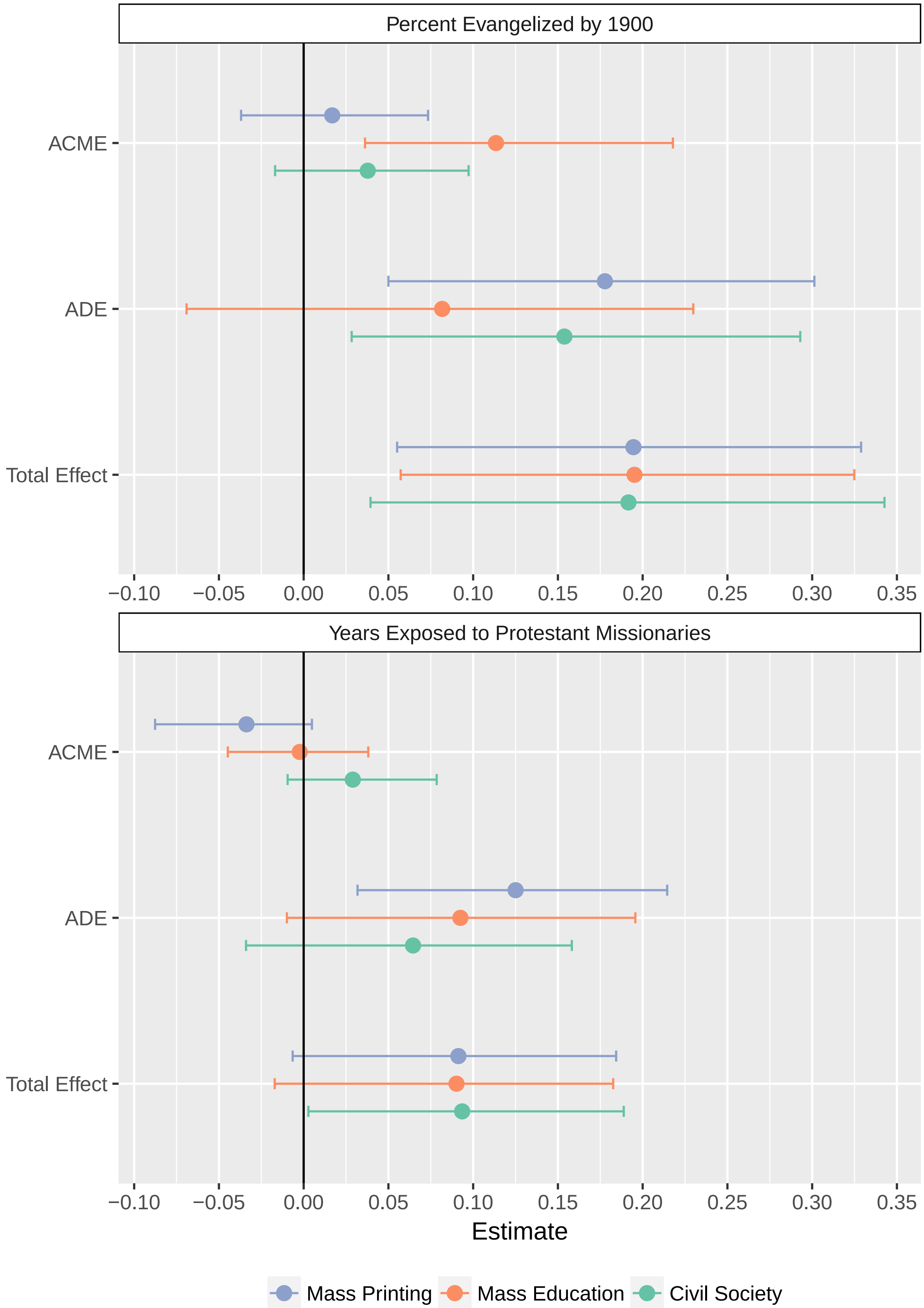
Note: The graphs present, for each mechanism, the sensitivity of the estimated ACMEs and ADEs to hidden confounding from an omitted pretreatment covariate. The dashed lines indicate the estimates assuming no hidden confounder.

Figure A8: ACME Sensitivity Analysis of the No Treatment-Mediator Interaction Assumption



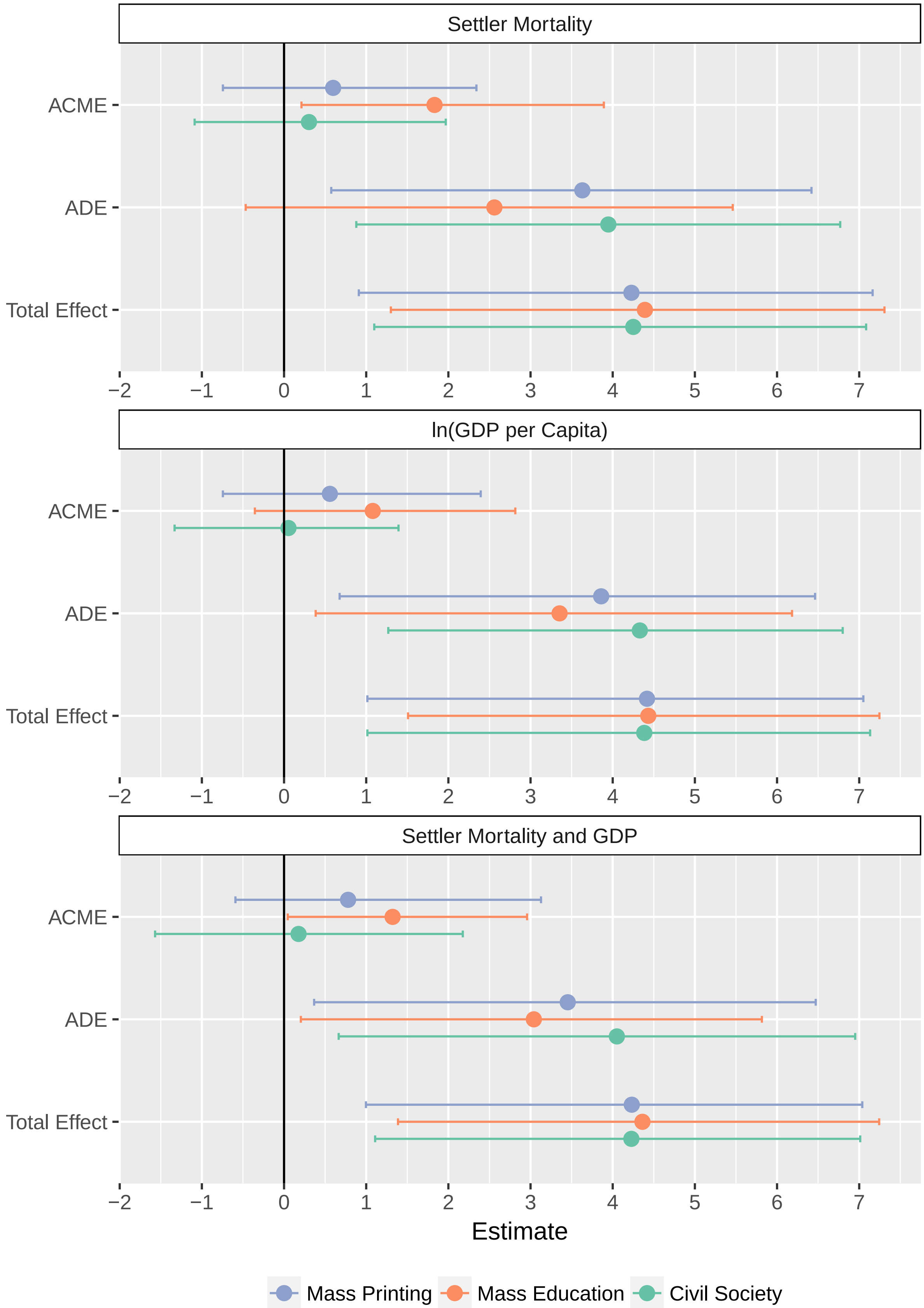
Note: The graphs present, for each mechanism, the sensitivity of the estimated ACMEs to increases in the variance explained by interaction between the treatment and mediator. The x-axes plot variance explained by the interaction and the y-axes plot the estimated ACME. The dashed lines indicate the estimates assuming no interaction.

Figure A9: Mediation Results with Alternative Treatment Variables



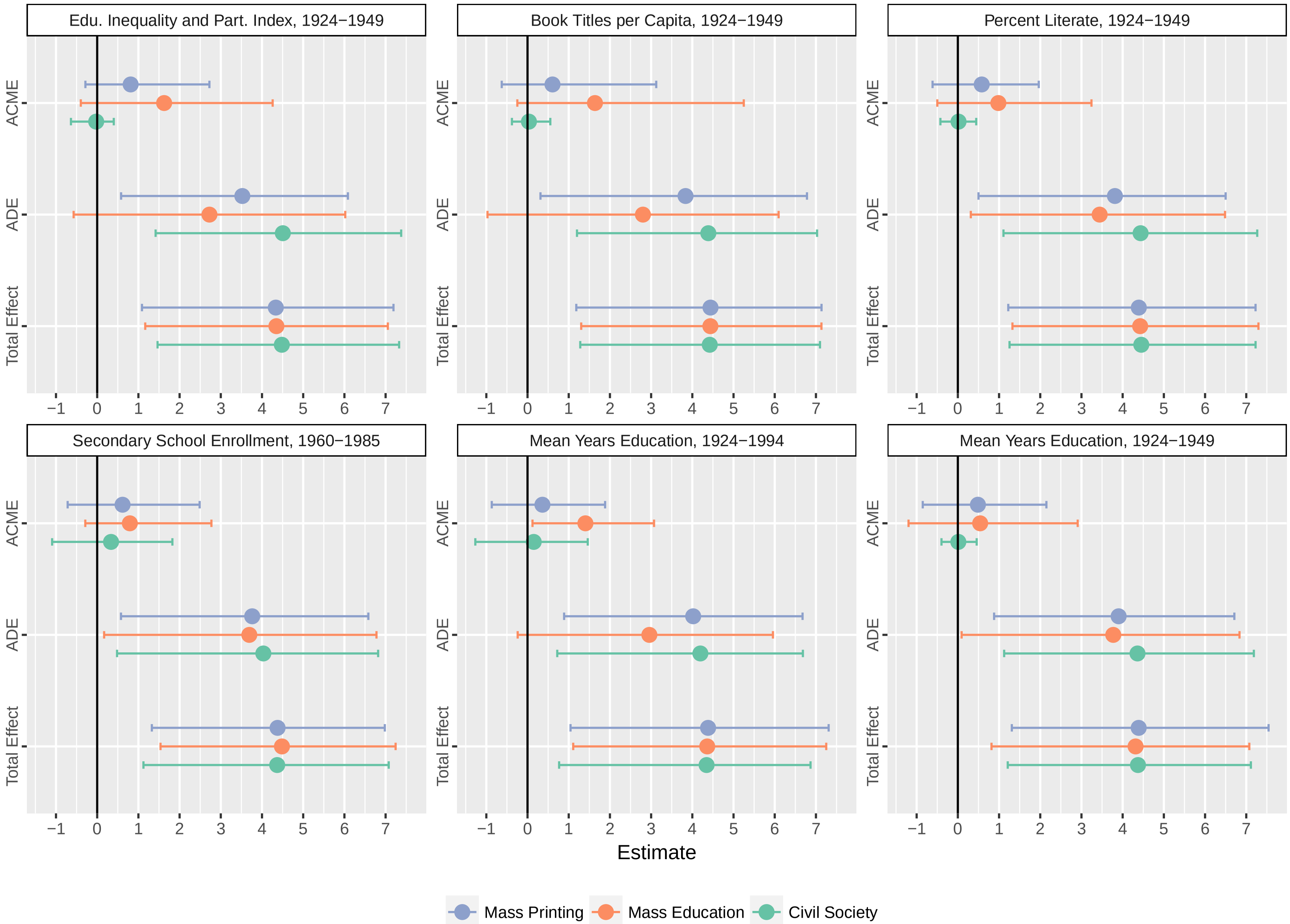
Note: The graphs present the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals.

Figure A10: Mediation Results with Additional Covariates



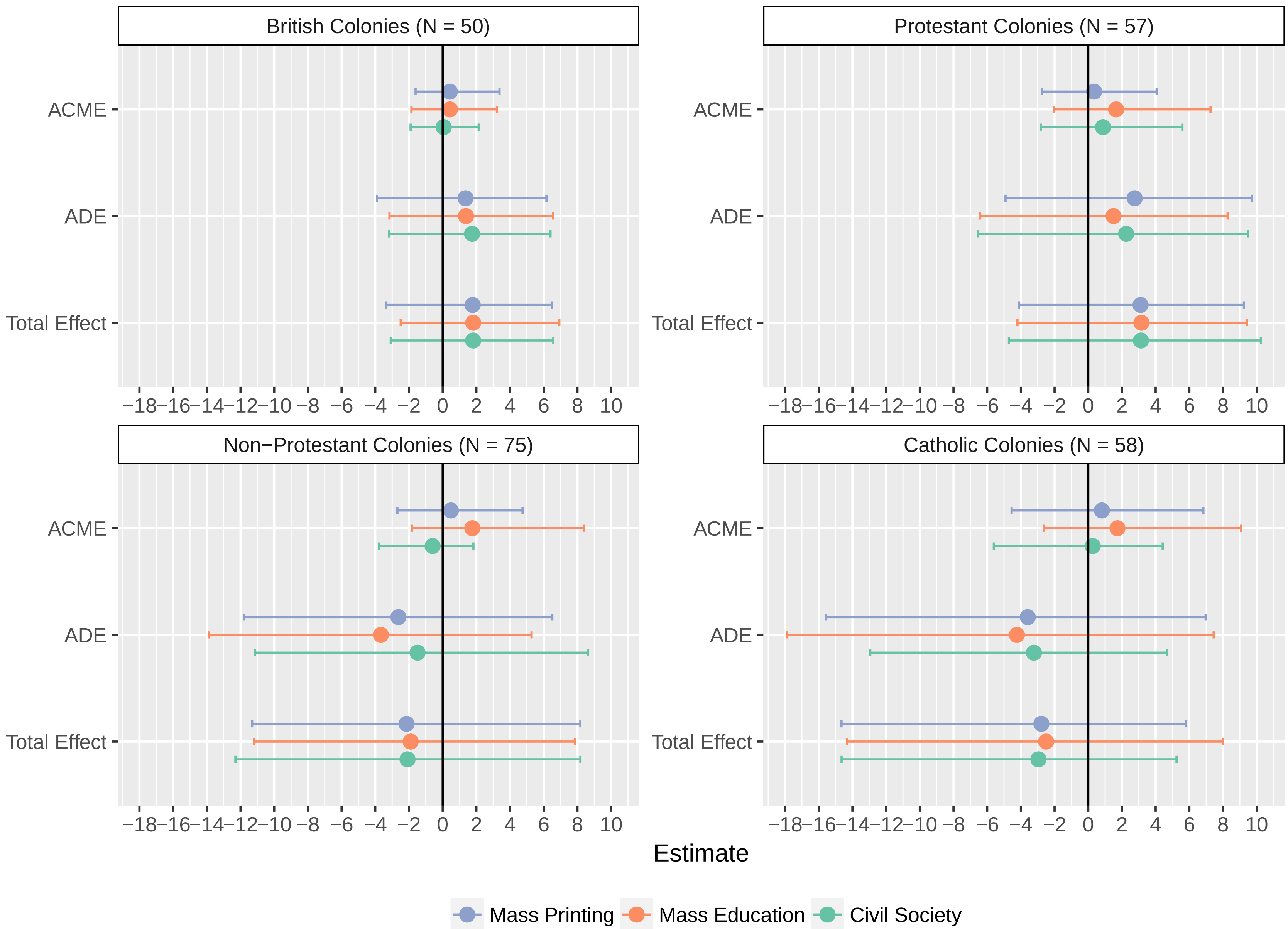
Note: The graphs present the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals.

Figure A11: Mediation Results with Alternative Mediator Measures



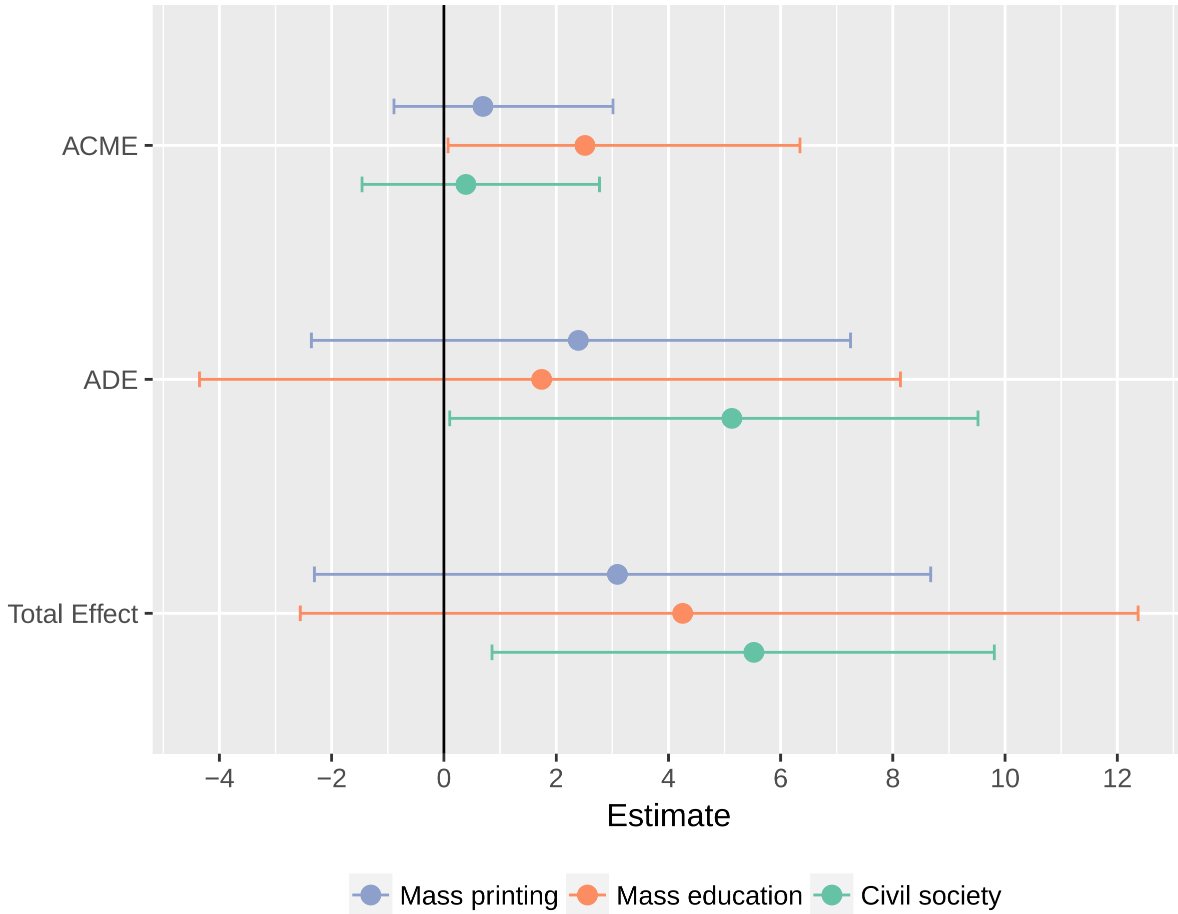
Note: The graphs present the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals.

Figure A12: Mediation Results Subset by Colonizer Types



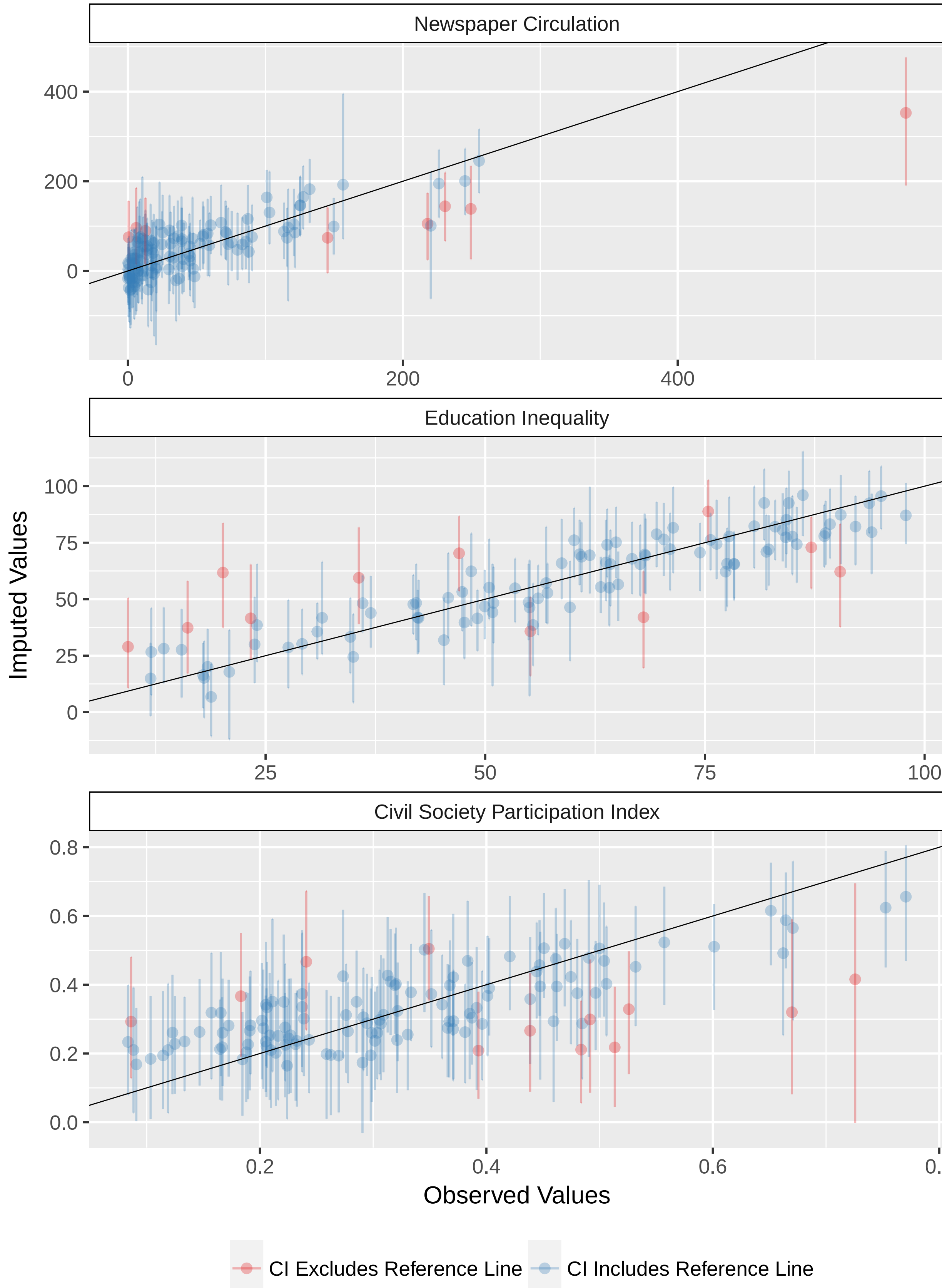
Note: The graphs present the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals.

Figure A13: Mediation Results without Using Multiple Imputation



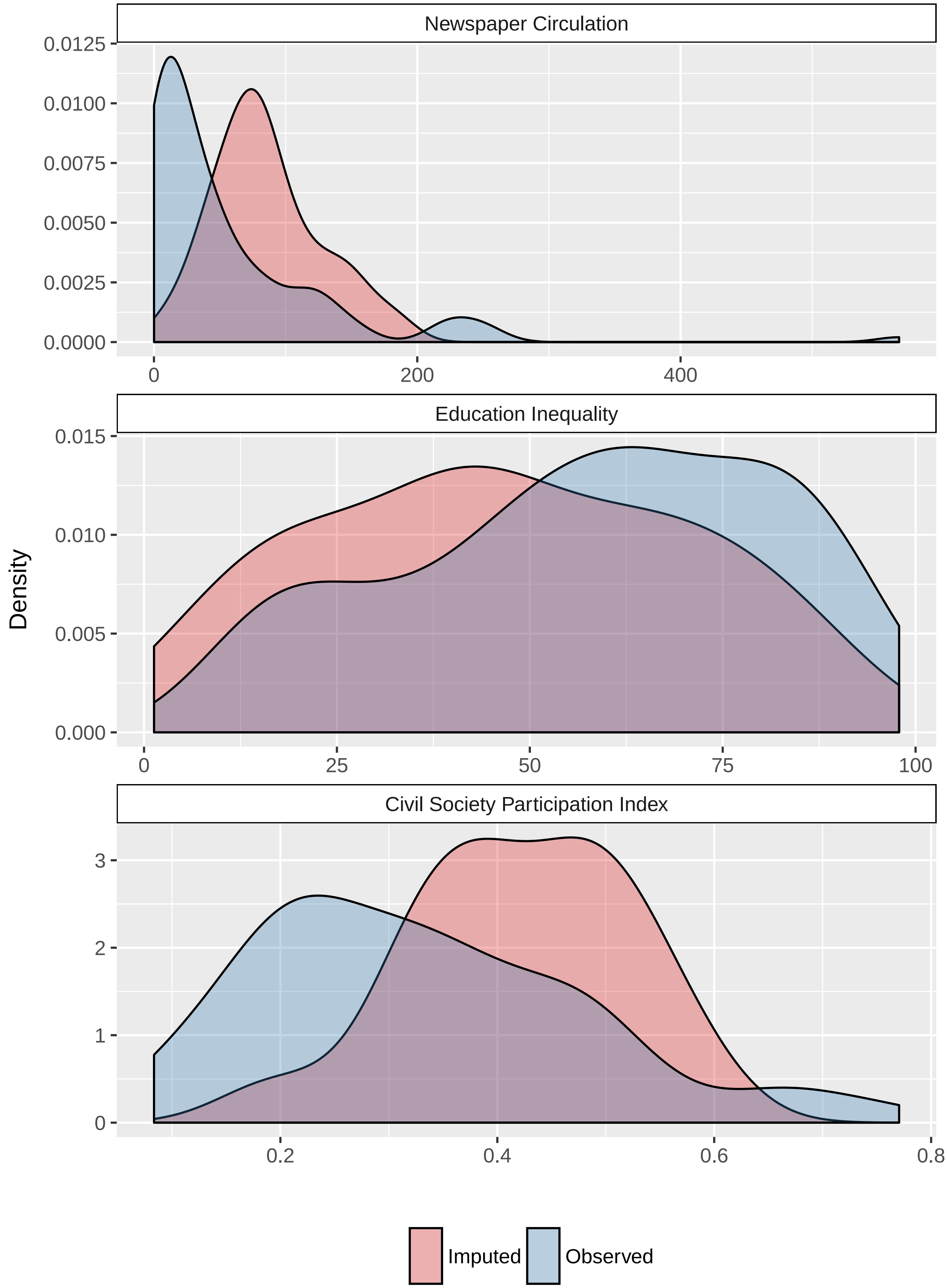
Note: The graph presents the estimated ACME, ADE, and total effects for each mediator. Line segments indicate 95% confidence intervals.

Figure A14: Overimputation Results for the Main Mediator Measures



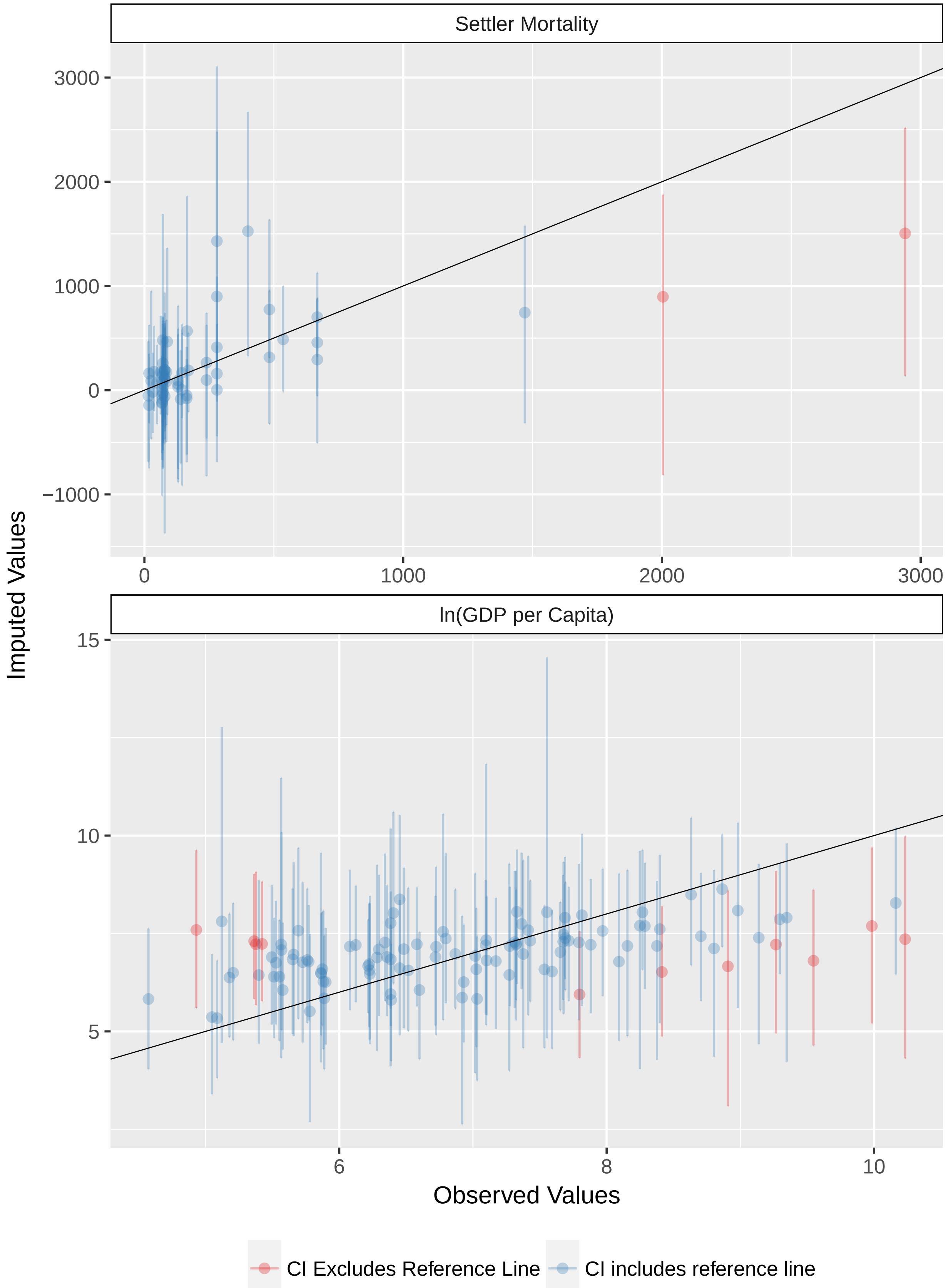
Note: The graphs present observed values of each mediator on the x-axes against mean imputations of those values on the y-axes. Line segments indicate 95% confidence intervals. The solid line serves as a reference point for perfect imputation.

Figure A15: Observed and Imputed Densities for the Main Mediator Measures



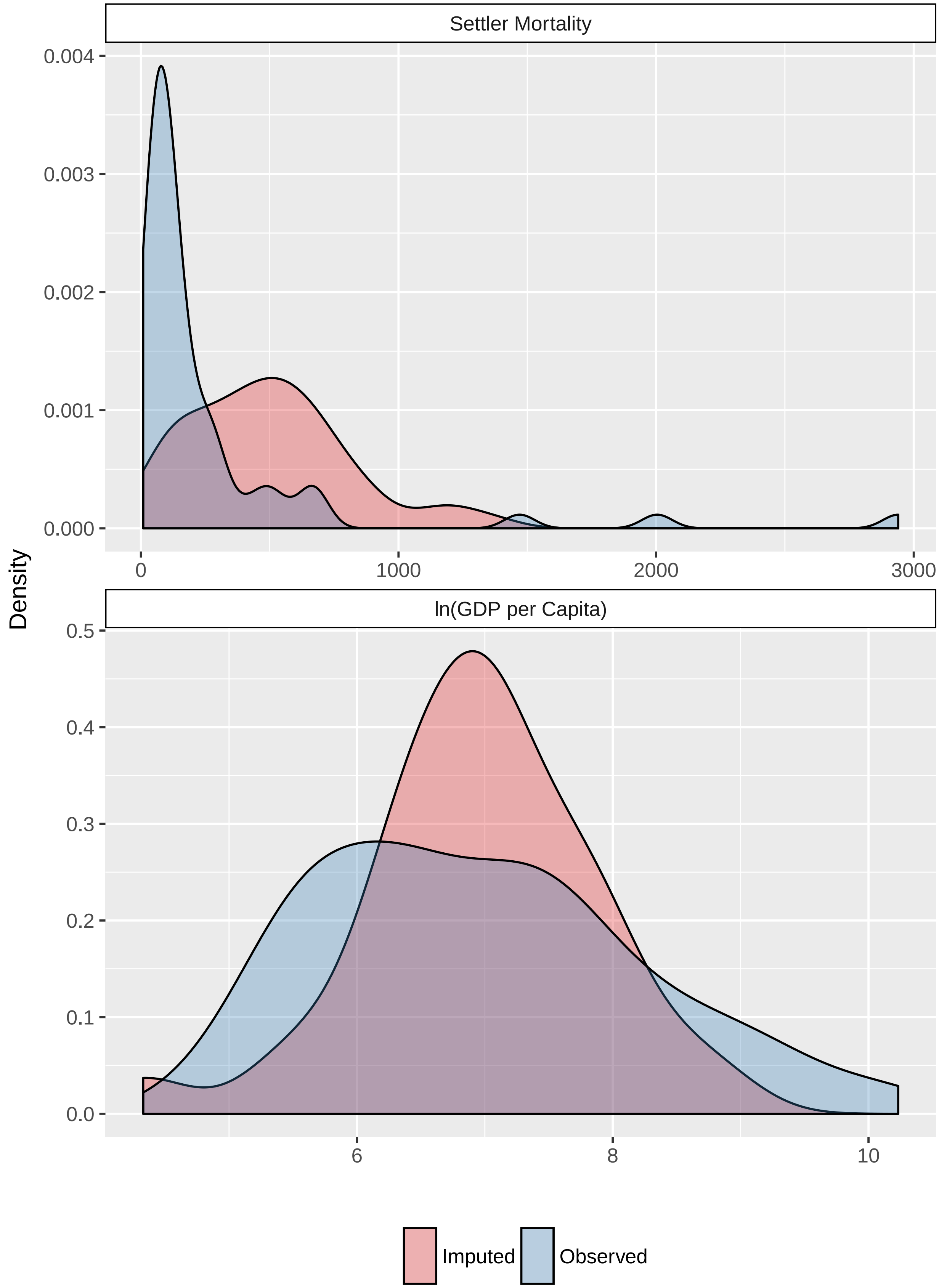
Note: The graphs present density plots of the observed and mean imputed values for each mediator.

Figure A16: Overimputation Results for Settler Mortality and GDP



Note: The graphs present observed values of each variable on the x-axes against mean imputations of those values on the y-axes. Line segments indicate 95% confidence intervals. The solid line serves as a reference point for perfect imputation.

Figure A17: Observed and Imputed Densities for Settler Mortality and GDP



Note: The graphs present density plots of the observed and mean imputed values for each variable.

Missionaries, Mechanisms, and Democracy

Preanalysis Plan

Kevin Angell[*] Jeffrey J. Harden[†]

Original version: June 7, 2018
This version: June 28, 2018
(see section 4 for updates)

Contents

1 Introduction

This document outlines a research plan for studying the causal mechanisms that characterize the effect of Conversionary Protestants (CPs) on the spread of liberal democracy. In an award-winning article published in the *American Political Science Review*, Woodberry (2012) demonstrates that CPs exerted a positive effect on democracy in the non-Western world. His identification

[*]Undergraduate Student, Department of Political Science, University of Notre Dame, 2060 Jenkins Nanovic Halls, Notre Dame, IN 46556, kangell@nd.edu.

[†]Associate Professor, Department of Political Science, University of Notre Dame, 2055 Jenkins Nanovic Halls, Notre Dame, IN 46556, jeff.harden@nd.edu.

strategy centers on two components: (1) an in-depth examination of the historical record with a focus on ruling out alternative explanations and (2) a quantitative analysis of data on 142 countries, which includes a comprehensive suite of regression specifications showing that the positive association between CPs and democracy is quite robust. Importantly, Woodberry (2012) provides a great deal of discussion about the causal mechanisms underlying this relationship. However, he notes that his empirical analyses "demonstrate a causal association between Protestant missions and democracy, but do not test which mechanism is most important" (256).

In this research, we plan to pick up where Woodberry (2012) left off. We will build on his work by assessing the three central causal mechanism that he posits in his discussion of the historical record: (1) mass printing, (2) mass education, and (3) civil society. Specifically, we will use causal mediation analysis (Imai, Keele, Tingley, and Yamamoto 2011; Imai and Yamamoto 2013) to evaluate the relative importance of each factor, as defined by their estimated mediation effects and proportions of the causal effect that flow through each one. These three mechanisms represent the core of Woodberry's (2012) theoretical contribution. Thus, examining their relative importance is a critical component of a comprehensive test of his theory. Below we provide details on the causal model we plan to evaluate and a research design we intend to execute. We also discuss a plan of action for incorporating additional analyses that we do not yet foresee into our work.

2 Causal Model

Woodberry (2012, 256) summarizes the causal process between CPs and democracy in a stylized causal graph, which we reproduce as Figure 1. The graph depicts an intricate system that includes pretreatment covariates, the treatment itself, several mediating variables (i.e., mechanisms), and ultimately the outcome. The large number of causal arrows indicates a complex set of dependencies between the relevant factors, which is not surprising given the complexity of the topic. Indeed, in his review of the literature Woodberry (2012) discusses an enormous volume of research addressing the spread of liberal democracy. However, there is an inherent tradeoff between the theoretical process depicted in that graph and what is feasible with the empirical evidence. Thus, the causal process as represented by Woodberry's (2012) statistical models is somewhat simpler.

[Insert Figure 1 here]

We intend to analyze a simpler version of this model in our mediation analysis as well. While we plan to use the same set of pretreatment covariates as Woodberry uses (see below), we will focus our attention on the three causal mechanisms to which he emphasizes the most and devotes the most attention to in his discussion of the historical record: mass printing, mass education, and civil society (Woodberry 2012, 249–253). These are not the only mediating variables in Figure 1. However, they stand out as the most theoretically-motivated mechanisms among the various factors he considers. In fact, Woodberry largely discusses the other mediating variables drawn in Figure 1 within the context of these main three mechanisms. Thus, while the mediation process may actually be more complex, mass printing, mass education, and civil society represent a useful group of variables on which to focus as we seek to design a feasible study of causal mechanisms.

Figure 2 presents two causal graphs of the mediation process we intend to test. First consider panel (a). The treatment is Conversionary Protestants and the outcome is Democracy. Woodberry's (2012) data are observational, and so it is necessary to include pretreatment covariates to mitigate confounding. These covariates may affect the treatment, the mediators, and/or the outcome directly without threatening identification of mediation effects (Keele, Tingley, and Yamamoto 2015). The three mediating variables—mass printing, mass education, and civil society—fall between the treatment and outcome. Additionally, we draw an arrow between the treatment and the outcome to allow for the possibility of a direct effect.

[Insert Figure 2 here]

Figure 2, panel (a) represents a causal process in which we assume no association between the three mechanisms. This choice makes estimation more straightforward. However, that assumption may not always be realistic (Imai and Yamamoto 2013). One possible violation in this case might be that increased access to printed materials among the mass public in a given country caused an increase in mass education. Woodberry (2012, 251) actually makes the point that such an association does *not* exist in the historical record, so our simpler model in panel (a) may be appropriate.

Nonetheless, we also plan to consider a more complicated process—shown in panel (b)—in which we relax the assumption of no association between printing and education. Throughout our analyses we maintain that civil society is distinct from the other mechanisms, and thus we assume its independence, conditional on the covariates.[1]

3 Research Design

Here we describe our empirical strategy for testing the causal models presented in Figure 2. In brief, our plan is to start with Woodberry's (2012) original regression analyses, then add the mediation analysis. We will use the same variables, coding decisions, estimation methods, and other choices as described in Woodberry (2012).[2] We adopt this approach because the primary goal of this research is to build on Woodberry's work. Moreover, we want our results to be as comparable as possible to Woodberry's results. Thus, we define the set of choices Woodberry made as the status quo.

As of the original deposit date of this document (June 7, 2018), we have acquired Woodberry's (2012) replication data and successfully replicated the exact results he presents in the article. We have also collected additional data and made choices about the new analyses we plan to conduct, which we discuss below. We have *not* yet conducted any mediation analyses. Thus, our plan for conducting this research is results-blind.

3.1 Data

Woodberry's (2012) outcome variable is a country's mean democracy score during the period 1950–1994, scaled from 0–100 (with higher scores indicating higher levels of democracy). The data come from the Cross-national Indicators of Liberal Democracy series (see Paxton 2002; Bollen 2009). Woodberry (2012) selects this measure because it includes a broader range of countries than alternatives and minimizes rater bias (Woodberry 2012, 257).

[1] We will empirically assess the between-mediator associations for all three mechanisms and the consequences of those associations for our mediation estimates. See our discussion of robustness checks below.

[2] We will report any deviations from this approach in our final manuscript and/or in updates to this document.

3.1.1 Pretreatment Covariates and Treatment

Woodberry (2012) collects several pretreatment covariates to include in his regression models.[3] Briefly, these covariates account for alternative theories of the spread of democracy, exogenous and precolonial conditions, other factors that influenced colonizers and missionaries, and endogenous or intervening variables (see Woodberry 2012, 257–258). Finally, he includes a set of "mission variables." Some of these variables are covariates, such as foreign Catholic priests per 10,000 population in 1923 and years exposure to Catholic missions. The remaining measures in this group relate to Protestant missions and comprise the set of variables of primary theoretical interest: percent Evangelized by 1900, years exposure to Protestant missions, and Protestant missionaries per 10,000 population in 1923.

While we include all of the mission variables in our analyses, we use this latter variable, Protestant missionaries per 10,000 population in 1923, as our treatment. It is preferable to the other two candidates (percent Evangelized and years exposure) because it is the most direct measure of the presence of CPs in a country. The percent Evangelized measure includes converts to Catholicism in it (Woodberry 2012, 257) and years exposure measures the earliest time at which CPs were present in a country rather than levels of CPs in that country (Woodberry 2012, 263).[4]

3.1.2 Mediators

Our mediator variables are designed to test the three causal mechanisms discussed above. We rely on a combination of data collected by Woodberry and our own data collection for these variables. These efforts yielded multiple candidates for each mechanism. We plan to complete analyses with all of them, as we describe here.

The first mediator is mass printing, which Woodberry (2012) primarily discusses as enhancement to the public sphere via the expansion of newspapers (249). Although he does not employ it

[3]Some of the covariates are not actually pretreatment as defined by measurement. For example, the treatment variable we select (see below) reflects measurement in 1923, but percent European is measured in 1980 and percent Muslim in 1970. While posttreatment bias is a serious one (Montgomery, Nyhan, and Torres 2018), we choose to bracket this issue entirely because the inclusion of these variables was a choice made by Woodberry. We place a very high premium on avoiding any deviations from the original specifications except for the addition of our mediation analysis.

[4]Specifically, Woodberry (2012) measures this variable as $1960 - y_c$, where y_c is the first year in which CPs arrived in country c (263).

in his main analyses, Woodberry's data include a measure of this conceptualization: an indicator of average daily newspaper circulation per 1,000 population in 1975, 1980,1985, and 1990.[5] We compute the means across these four years to construct a mass printing mediator variable. The measure itself is internally valid; however, the timing is not ideal. Measurement does occur post-treatment (i.e., after 1923), but it also falls after the measurement of the outcome begins (1950). Accordingly, we also plan to consider an indicator that can be measured earlier: the number of book titles per capita (Fink-Jensen 2015).[6] This variable is useful in that it can be measured in the 1924–1949 period that separates treatment and outcome. However, the data have a much larger amount of missingness for Woodberry's sample of countries than Woodberry's newspaper data (see Table 1 and discussion below).

Next, we consider mass education as a mediating variable. Woodberry's data also contain a candidate measure, which he uses in some robustness checks (Woodberry 2012, Table 5, 265). Specifically, his data include the mean enrollment in secondary education from 1960–1985 (Barro and Lee 1994). However, this variable does not cover all of the countries in his full sample and is also partially concurrent with the democracy outcome. As an alternative, we gathered education data from the Varieties of Democracy Project (V-Dem, see Coppedge et al. 2018). V-Dem provides data on the percentage of a country that is literate (from Vanhanen 2003), the average years of education for citizens older than 15, and a Gini-type measure of educational inequality. Each of these variables are available at least as early as 1924.

We consider the percentage literate variable to be a particularly promising mediator variable for both theoretical and empirical reasons. Theoretically, it matches Woodberry's (2012) argument that CPs sought to provide education to spread literacy so that individuals could read the Bible themselves (246, 249). This literacy, in turn, promoted democracy (Woodberry 2012, 251). Empirically, the literacy data cover more countries than Woodberry's secondary education enrollment data. The average years of education variable allows a more detailed measurement of education levels compared to Woodberry's measure, particularly for countries with relatively few individuals

[5]These variables come from the United Nations (UN) Data (see http://data.un.org/).

[6]This measure is available at https://www.clio-infra.eu/Indicators/BookTitlesperCapita.html.

who receive secondary education. However, this variable suffers from a large amount of missing data for Woodberry's sample of countries when compared to the literacy variable. Finally, the educational inequality variable allows us to examine a slightly different aspect of mass education that Woodberry proposes. He argues that CPs spread education to non-elites, reducing educational inequality (Woodberry 2012, 251). However, this variable also suffers from more missing data than the literacy data.

Finally, civil society is the third causal mechanism we consider. Although the development and spread of voluntary organizations, nonviolent protest, and other political movements is a key element of Woodberry's (2012) theoretical framework (e.g., 252–253), he does not include any such measure in his analyses or replication data. To obtain one, we again turn to V-Dem (Coppedge et al. 2018). V-Dem provides a civil society participation index that captures the extent to which citizens are involved in CSOs, how much CSOs are consulted by policy makers, whether women participate in CSOs, and whether political candidate nomination is decentralized (Coppedge et al. 2018). This variable is highly suitable to serve as a mediator because it provides a robust measurement of the civil society that Woodberry claims connects CPs with the development of democracy and covers the vast majority of the countries in the estimation sample (Woodberry 2012, 252–253).

As our discussion above makes clear, the key issue that we must address with all of these possible measures of causal mechanisms is missing data. Our objective is to use Woodberry's full sample of data in our mediation analyses. His measurement strategy involves computing several variables by averaging over a span of years; for instance, his outcome variable is a mean over the period 1950–1994. We adopt this approach in constructing our mediators. However, the historical nature of the data means that some countries did not exist or were known by different names when our various mediators were measured.[7] Consequently, we consider two options when constructing mediators: compute the averages from (1) 1924–1949 or (2) 1924–1994. The former approach best matches the temporal nature of a mediator because it occurs between treatment (1923) and

[7]See Woodberry (2012, 257) for a discussion of how he addresses this issue in measuring the outcome variable.

outcome (1950). However, it yields more missing data. In contrast, the second strategy overlaps measurement of the outcome, but allows for more data to be collected. We plan to conduct our analyses using both measurement strategies to assess the robustness of our results. Table 1 summarizes our mediator data collection efforts.

[Insert Table 1 here]

For those data values we cannot fill in, we plan to use multiple imputation with the Amelia II software available in R (Honaker, King, and Blackwell 2011). Amelia begins with the assumption that the complete data are distributed multivariate normal, then employs the expectation-maximization (EM) algorithm to impute the missing values. It repeats this process m times (we plan to use the default of $m = 5$), then analysis continues as usual with those m complete datasets. An adjustment to measures of uncertainty is also necessary (see Blackwell, Honaker, and King 2017, 309).[8] We will use the complete data from Woodberry's sample to impute the missing values in our mediators.[9] Amelia also includes diagnostic functions for checking the imputation results. We plan to use these tools to ensure the validity of our imputed mediators before proceeding to our main analyses.

3.2 Specification Selection

Woodberry's (2012) empirical analysis is comprised of 32 reported regression models, with many more in the replication files. For purposes of feasibility, we must simplify before proceeding to the mediation analysis. Based on its comprehensive nature and Woodberry's (2012) discussion of the empirical results, we regard his Model 4 from Table 3 (262) as the "Canonical Model."[10]

[8]We will perform this adjustment according to the following steps, as discussed by Blackwell, Honaker, and King (2017, 309). First, we will simulate our quantities of interest from the mediation models with each of the m datasets via bootstrapping and/or quasi-Bayesian replicates (see Tingley, Yamamoto, Hirose, Keele, and Imai 2014). Second, we will combine the simulation replicates as if they came from the same model. Finally we will summarize the combined vectors of simulated replicates to report the results (see also King, Tomz, and Wittenberg 2000; Imai, King, and Lau 2008).

[9]We do not need to impute any data from the covariates that Woodberry (2012) includes in his regressions. We only need to impute data in the mediators.

[10]Another candidate model is Table 2, Model 3 (Woodberry 2012, 260). We favor Table 3, Model 4 because it includes covariates measuring the process of colonization (Woodberry 2012, 262–263). However, we plan to assess the robustness of our results using Table 2, Model 3 (see below).

This model is estimated on his full sample of 142 countries and includes several groups of covariates (though not every single covariate available). Moreover, among the models he reports that are estimated on the full sample, this model produces the second-largest adjusted-R^2 value.[11] Thus, it is a good representation of the statistical and substantive results that Woodberry communicates in his research. We focus our analyses on this model.

The Canonical Model reports a positive and statistically significant treatment effect of 4.43 (confidence interval: [1.28, 7.59]). Recall that the outcome variable is scaled from 0–100. The estimate indicates that, all else equal, a standard deviation increase in Protestant missionaries is associated with a 0.26 standard deviation increase in a country's average democracy score. Woodberry highlights the substantive significance of this result throughout his discussion. In brief, while there are many factors that affected the spread of democracy, the effect of CPs is strong enough that it deserves serious scholarly attention.

3.3 Estimation

Woodberry (2012) employs various linear regression estimators in his analyses, including ordinary least squares (OLS), robust regression (RR), and instrumental variables (IV). He estimates the Canonical Model with RR due to concerns with skewness and outliers in some of the covariates (Woodberry 2012, 259). This estimator uses iteratively reweighted least squares (IRLS) to weight observations based on their outlyingness (e.g., Street, Carroll, and Ruppert 1988). We plan to use Stata's implementation of the estimator as Woodberry (2012) did to maintain comparability to the original results. We plan to conduct the mediation analysis using the mediation package in R (Tingley et al. 2014).[12]

[11] The adjusted-R^2 reported in Table 3, Model 5 is somewhat larger (0.496 versus 0.467). However, Model 5 is incorrectly specified because it omits some constituent terms of interaction effects (Brambor, Clark, and Golder 2006).

[12] We can use Stata's implementation of RR in R by estimating the model in Stata and saving the weights generated by the estimator, then using those weights in a standard OLS estimation in R. This approach reproduces the coefficients from Stata exactly, although the standard errors are off. We plan to employ bootstrapping and/or robust standard errors in the mediation analysis to correct this issue.

3.3.1 Main Analyses

We plan to focus on two main analyses. First, we will estimate the causal mediation effect of each mediator separately. This approach is consistent with the causal graph in Figure 2, panel (a), in which we assume no association between the mediators. Our key quantities of interest in this analysis will be the average causal mediation effect (ACME), which is the effect that flows through the mediator, and the proportion mediated (PM), which indicates the relative size of the ACME to the total effect (Tingley et al. 2014). We plan to (1) compare PM values for each mediator and (2) formally test the null hypothesis that the three ACMEs are equal to each other using the $\alpha = 0.05$ threshold for statistical significance.

Second, we will repeat the first analysis, but allow for the possibility of association between mass printing and mass education. This association is reflected in the causal graph depicted in Figure 2, panel (b). To complete this analysis we will use the sensitivity analysis approach recommended by Imai and Yamamoto (2013). The key quantities of interest will remain the same as in the first analysis.

3.3.2 Additional Robustness Checks

In addition to our main analyses, we also plan to conduct the following robustness checks and alternative specifications. We will most likely report these results in an appendix to our final manuscript.

- **Second model specification**. We will repeat our main analyses using the specification in Woodberry's (2012) Table 2, Model 3 (260).
- **Alternative treatment variables**. We will repeat our main analyses with Woodberry's other primary variables of interest—percent Evangelized by 1900 and years exposure to Protestant missions—as the treatment variable.
- **Additional covariates**. We will repeat our main analyses after including two covariates that Woodberry includes in other specifications, but does not include in the Canonical Model: economic development (operationalized as logged gross domestic product [GDP] per capita) and settler mortality (Acemoglu, Johnson, and Robinson 2001).

- **Additional mediator measures**. We will repeat our main analyses with all of the measures of each mediator described above.
- **Subsetting by region**. We will repeat our main analyses in several regions of the world discussed by Woodberry (Africa, Latin America, Asia, and Oceania).
- **Subsetting by colonizer**. We will repeat our main analyses after grouping countries by colonizer according to Woodberry's definitions (British, Catholic, other religious liberty colonies, Dutch, not colonized).
- **Omitting multiple imputation**. We will reduce the sample to only the cases for which we have mediator data without using multiple imputation and repeat the main analyses.
- **Sensitivity to covariates**. We will conduct sensitivity analyses on the possibility of an unmeasured confounding pretreatment covariate (Imai et al. 2011).
- **Sequential ignorability**. We will conduct sensitivity analyses that allow each mediator to be associated with the other mediators (Imai and Yamamoto 2013).

4 Updates

The original version of this plan was deposited on June 7, 2018. We plan to update it as needed over the course of the project. If we make any changes to our analysis after depositing this document, we will explain and justify those changes here.

4.1 June 28, 2018: Preferred Model Specification

In working with the data, we have identified potential specification issues in the Canonical Model that we would like to address. Accordingly, we plan to conduct a supplemental mediation analysis on our own "preferred specification" rather than any particular model reported in Woodberry (2012). Our preferred specification begins with the Canonical Model and addresses several key issues: posttreatment bias, treatment effect heterogeneity, treatment definition, and degrees of freedom.

First, we remove all of the covariates that are measured after 1923 to avoid posttreatment bias (Montgomery, Nyhan, and Torres 2018). These include percent European and percent Muslim, as well as the additional covariate of GDP per capita (logged), which is not in the Canonical Model but is included in one of our preregistered robustness checks. All of these variables are measured in the latter half of the 20th Century. We also include settler mortality (and impute its missing values) due to its theoretical relevance (Woodberry 2012, 263). This variable was measured in the mid-1800s (see Acemoglu, Johnson, and Robinson 2001, 1382).

Next, we include indicator variables for five regions of the world: Sub-Saharan Africa, Asia, Latin America and the Caribbean, the Middle East/North Africa, and Oceania (the reference category). We plan to interact these region indicators with the treatment variable to assess variation in the treatment effect by region.

We also reduce down to a single treatment variable for the sake of definition clarity: Protestant missionaries per 10,000 population in 1923. Consistent with the discussion from above about the three variables Woodberry (2012) employs, we omit percent Evangelized by 1900 and years exposure to Protestant missions. We retain foreign Catholic priests per 10,000 population in 1923 and years exposure to Catholic missions as covariates.

Finally, we simplify the model by removing several variables associated with the perceived value of a country. While Woodberry (2012) provides theory for these variables, only one—an indicator for whether a Protestant Colonizer took a colony from a Catholic Colonizer—yields a substantively and statistically significant coefficient estimate. Thus, to save degrees of freedom, we only retain that variable in our preferred specification.

The variables included in our preferred model specification are listed below as follows. Covariates in bold are included in our preferred model and the Canonical Model. Covariates in plain type are only included in the Canonical Model (i.e., dropped in our preferred specification). Finally, italics represents variables that are not included in the Canonical Model, but included in our preferred specification.

- **British Colony**
- **Other Religious Liberty Colony**
- **Dutch Colony**
- **Never Colonized Significantly**

- **Latitude**
- **Island Nation**
- **Landlocked Nation**
- Percent European in 1980
- Percent Muslim in 1970
- **Major Oil Producer**
- **Literate Culture Before Missionary Contact**
- Years Exposure to Protestant Missions
- **Protestant Missionaries per 10,000 population in 1923 (Treatment)**
- Percent Evangelized by 1900
- **Years Exposure to Catholic Missions**
- **Foreign Catholic Priests per 10,000 population in 1923**
- **Year of 1st Democracy Data**
- **Post-1976 Democracy Data Only**
- Date 1st Sighted by Europeans after 1444
- Gap between Sighted and 1st Missionaries
- Mission Gap × Literacy
- Mission Gap × Latitude
- Gap between Sighted and Colonized
- Colonial Gap × Literacy
- Colonial Gap × Latitude
- Number of Times Territory Switched Colonizers
- **Protestant Colonizer Took Colony from Catholics**
- *Settler Mortality*
- *Region Indicator Variables (Sub-Saharan Africa, Asia, Latin America and the Caribbean, the Middle East/North Africa, and Oceania)*
- *Region Indicators × Treatment*

We plan to conduct moderated mediation analysis with this specification to assess whether the mediation effects vary by region.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development." *American Economic Review* 91(5): 1369–1401.

Barro, Robert J., and Jong-Wha Lee. 1994. "Data Set for a Panel of 138 Countries." http://www.nber.org/pub/barro.lee/.

Blackwell, Matt, James Honaker, and Gary King. 2017. "A Unified Approach to Measurement Error and Missing Data: Overview and Applications." *Sociological Methods & Research* 46(3): 303–341.

Bollen, Kenneth A. 2009. "Liberal Democracy Series I, 1972–1988." *Electoral Studies* 28(3): 368–374.

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14(1): 63–82.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jerey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig, and Daniel Ziblatt. 2018. "V-Dem Codebook v8." Varieties of Democracy (V-Dem) Project. https://www.v-dem.net/.

Fink-Jensen, Jonathan. 2015. "Book Titles per Capita." IISH Dataverse. http://hdl.handle.net/10622/AOQMAZ.

Honaker, James, Gary King, and Matthew Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45(7): 1–47.

Imai, Kosuke, and Teppei Yamamoto. 2013. "Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments." *Political Analysis* 21(2): 141–171.

Imai, Kosuke, Gary King, and Olivia Lau. 2008. "Toward a Common Framework for Statistical Analysis and Development." *Journal of Computational Graphics and Statistics* 17(4): 1–22.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4): 765–789.

Keele, Luke, Dustin Tingley, and Teppei Yamamoto. 2015. "Identifying Mechanisms Behind Policy Interventions via Causal Mediation Analysis." *Journal of Policy Analysis and Management* 34(4): 937–963.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2): 341–355.

Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. "How Conditioning on Post-treatment Variables Can Ruin Your Experiment and What to Do about It." Forthcoming, *American Journal of Political Science*.

Paxton, Pamela. 2002. "Social Capital and Democracy: An Interdependent Relationship." *Ameri-*

can Sociological Review 67(2): 254–277.

Street, James O., Raymond J. Carroll, and David Ruppert. 1988. “A Note on Computing Robust Regression Estimates Via Iteratively Reweighted Least Squares.” *The American Statistician* 42(2): 152–154.

Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele, and Kosuke Imai. 2014. “mediation: R Package for Causal Mediation Analysis.” *Journal of Statistical Software* 59(5): 1–38.

Vanhanen, Tatu. 2003. *Democratization: A Comparative Analysis of 170 Countries*. New York: Routledge.

Woodberry, Robert D. 2012. “The Missionary Roots of Liberal Democracy.” *American Political Science Review* 106(2): 244–274.

Figure 1: Stylized Causal Graph of the Effect of Conversionary Protestants on the Spread of Liberal Democracy (Woodberry 2012, Figure 1, 256)

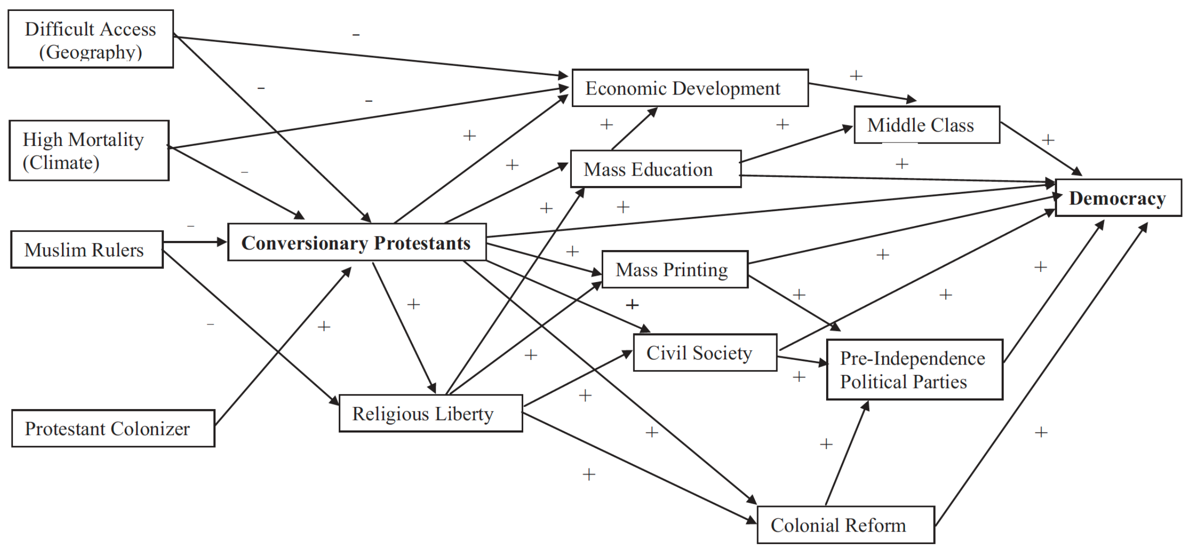
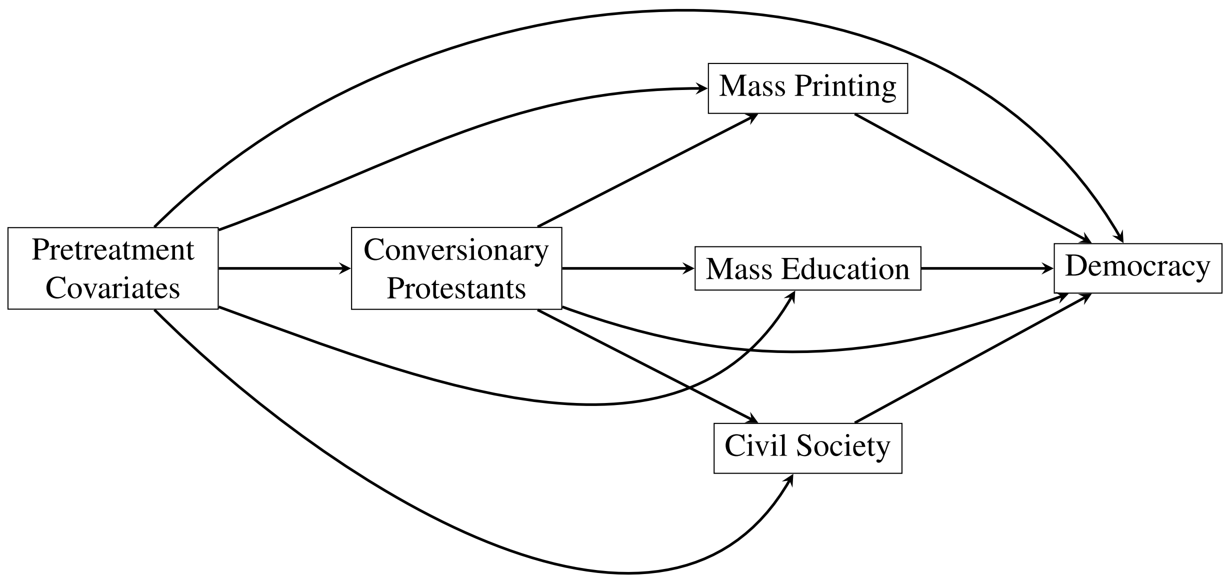


Figure 2: Proposed Causal Graph for Empirical Evaluation of the Mechanisms

(a) No Mediator Association



(b) Association Between Printing and Education

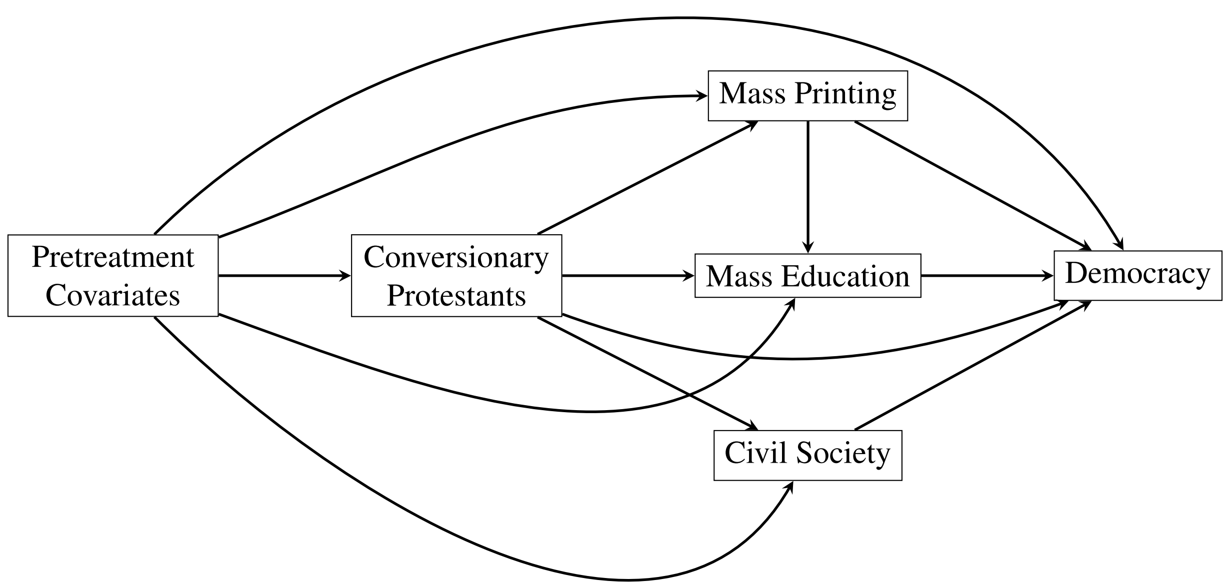


Table 1: Missingness in the Proposed Measures of Causal Mechanisms

Mediator	Variable	Source	Cases Missing	% Missing
Mass Printing	Mean daily newspaper circulation (average of 1975, 1980, 1985, 1990)	Woodberry (2012), UN Data	24	16.9
	Mean book titles per capita (average)	Fink-Jensen (2015)	129/112	90.8/78.9*
Mass Education	% Literate (average)	Vanhanen (2003); Coppedge et al. (2018)	96/32*	67.6/22.5*
	Mean years education (average)	Coppedge et al. (2018)	77/42*	54.2/29.6*
	Secondary school enrollment (1960–1985)	Woodberry (2012)	57	40.1
	Gini educational inequality (average)	Coppedge et al. (2018)	76/42*	53.5/29.6*
Civil Society	Civil society participation index (average)	Coppedge et al. (2018)	27/18*	19.0/12.7*

Note: Cell entries summarize missing data in the mediator measures. The percentage missing for each variable is computed based on the 142 countries in the estimation sample of Woodberry's (2012) Table 3, Model 4 (262). * Pairs of entries denoted ·/· indicate the two measurement strategies discussed in the text: computing the averages from 1924–1949 (listed first) or 1924–1994 (listed second).